

KULKUTAVANPÄÄTTELYALGORITMIT GPS-AINEISTOSTA

Annika Rantala

Aalto-yliopisto
Teknillinen korkeakoulu
Rakennus- ja ympäristötekniikan kirjasto

Teknillisen korkeakoulun yhdyskunta- ja
ympäristötekniikan laitoksella professori Timo
Ernvallin valvonnassa tehty diplomityö.

Espoo 14.12.2009

Tekijä:	Annika Rantala		
Diplomityö:	Kulutusalan päätelyalgoritmit GPS-aineistosta		
Päivämäärä:	14.12.2009	Sivumäärä:	95 + 32 liit.
Professuuri:	Liikennetekniikka	Koodi:	Yhd-71
Valvoja:	Prof. Timo Ernvall		
Ohjaaja:	DI, MBA Virpi Pastinen		
Avainsanat:	henkilöliikennetutkimus, liikkumistutkimus, GPS, kulutusalan päätely		

Tämän työn tarkoituksena oli selvittää, millaisia mahdollisuuksia GPS-paikannus tuo liikkumisen tutkimiseen. Työssä keskityttiin lähinnä henkilöliikennetutkimuksiin, ja niissä tarkemmin kulutusalan päätelyyn GPS-aineistosta. Tuloksia voidaan kuitenkin soveltaa myös muunlaisissa liikenne- ja liikkumistutkimuksissa, joissa kerätään tietoa muun muassa liikkumisen syistä, käytetyistä kulutusaloista ja matkojen ominaisuuksista.

Perinteisten tutkimusmenetelmien, kuten haastattelujen ja matkapäiväkirjojen, käytössä on omat rajoituksensa. Vastaaminen on usein työlästä, joten vastausprosentit pienenevät, osa matkoista unohdetaan ja arviot matkojen kestoista ja pituuksista ovat epätarkkoja. Tutkimusten avuksi on siksi kehitetty uusia menetelmiä. Niiden tarkoituksena on automatisoida tiedonkeruuta ja siten helpottaa tutkimuksiin osallistumista, parantaa kerätyn aineiston laatua ja parhaassa tapauksessa myös pienentää kustannuksia. GPS-laitteiden käyttö on yksi tällainen menetelmä.

Aluksi selvitettiin kirjallisuudesta, millaisia mahdollisuuksia ja toisaalta rajoituksia GPS-paikannus tuo liikkumistutkimuksiin ja millaisia tutkimuksia on jo toteutettu. Lisäksi selvitettiin, millaisia automatisoituja prosesseja tarvitaan, jotta GPS-aineistosta voidaan päätellä liikkumistutkimuksissa tarvittava tietoja, kuten kulutusala, matkan tarkoitus ja reitti. Erityisesti tutkittiin algoritmeja, joilla GPS-aineistosta voidaan päätellä käytetty kulutusala.

Työssä toteutettiin 50 hengen pilottitutkimus, jonka osallistujat keräsivät GPS-laitteilla aineistoa liikkumisestaan ja pitivät samalla tarkkaa matkapäiväkirjaa. Kerätyn aineiston perusteella toteutettiin kulutusalan päätelyalgoritmi.

Kulutusalan päätelyalgoritmien hyvyys riippuu siitä, mitä kulutusaloja on tarpeen erottaa ja mitä lähtötietoja on käytettävissä. Kirjallisuudessa esiintyvät parhaat algoritmit tunnistavat kulutusalan oikein jopa 95 prosentissa matkoista. Tässä työssä kehitetty algoritmi tunnisti kulutusalan oikein noin 80 prosentissa matkoista pääkaupunkiseudun oloissa. Tunnistettavat kulutusalat olivat kävely, pyöräily, auto, bussi, raitiovaunu, juna ja metro ja lähtötietoina käytettiin GPS-aineiston lisäksi joukkoliikenteen pysäkkien paikkoja. Päätelytulosta on mahdollista parantaa esimerkiksi käyttämällä muitakin paikkatietoja, kuten joukkoliikenteen reittejä.

GPS on lupaava menetelmä, mutta sillä ei pystytä keräämään kaikkia niitä tietoja, joita esimerkiksi henkilöliikennetutkimuksissa nykyisin kerätään. Siksi GPS ei voi täysin korvata nykyisiä tutkimusmenetelmiä, vaan täydentää niitä. GPS:n käyttö tuo tutkimuksiin myös uudenlaisia riskejä. On pystyttävä etukäteen arvioimaan kustannuksia ja kerättävän tiedon laatua.

Author:	Annika Rantala		
Thesis:	Mode Detection Algorithms for GPS Data		
Date:	14th December 2009	Number of pages:	95 + 32 app.
Professorship:	Transportation Engineering	Code:	Yhd-71
Supervisor:	Prof. Timo Ernvall		
Instructor:	M.Sc. (Tech), MBA Virpi Pastinen		
Keywords:	household travel survey, travel survey, GPS, mode detection		
<p>The purpose of this study was to find out what kinds of possibilities GPS has to offer for travel surveys. The study focused mainly on household travel surveys and, more specifically, the detection of mode of transport from GPS data. The results can also be applied to other kinds of traffic and travel surveys where the reasons behind travel, methods of transport and characteristics of trips are studied.</p> <p>Traditional methods, such as interviews and travel diaries, have their limitations. Answering all the questions is often a time-consuming task for respondents, response rates are falling, some trips are forgotten altogether, and estimates about travel distances and durations are inaccurate. Because of these weaknesses, new methods have been developed. Their goal is to automatize the data collection and thus facilitate participation in studies, enhance data quality and, if possible, also to reduce costs. Using GPS devices is but one of these new methods.</p> <p>First, relevant literature was examined to find out what kinds of possibilities GPS offers to travel surveys, what limitations there are to the technique and what kinds of surveys have already been carried out. In addition, automatized processes for detecting information needed in travel surveys – mode, purpose of trip, route – from GPS data were examined. The focus was especially on algorithms for detecting the mode of transport from GPS data.</p> <p>The study included a pilot study of 50 persons carrying a GPS device to gather travel data. They also kept an extensive travel diary including times and modes of transport used. Based on this data, a mode detection algorithm was developed.</p> <p>The quality of a mode detection algorithm depends on what modes need to be distinguished and the information available on the traffic system. The best algorithms can detect the mode correctly in 95 percent of trips. In this study, the result was around 80 percent. The modes detected were walking, bike, car, bus, tram, metro and train, and geographic information about public transport stops was used in addition to the GPS data. It is possible to improve the results, for example by using more geographic information such as public transport routes.</p> <p>Using GPS is a promising method, but it cannot be used to gather all the information that is nowadays gathered in household travel surveys. Thus GPS cannot fully replace the methods currently in use, but can only complement them. Using GPS also brings new risks to the surveys. The costs and the quality of data have to be assessed beforehand.</p>			

ALKUSANAT

Tämä diplomityö on tehty Teknillisen korkeakoulun yhdyskunta- ja ympäristötekniikan laitokselle. Työn tilaajina ja rahoittajina ovat olleet Pääkaupunkiseudun yhteistyövaltuuskunta YTV ja Tiehallinto. Käytännössä olen tehnyt diplomityöni WSP Finland Oy:ssä. Tiehallinnon rahoitus tuli älykkään liikenteen ohjelma ÄLListä, joten hieman muokattu raportti tästä työstä julkaistaan myös ÄLLi-julkaisusarjassa.

Työtä on valvonut professori Timo Ernvall Teknillisestä korkeakoulusta ja ohjannut Virpi Pastinen WSP Finland Oy:stä. Tilaajien puolelta työn ohjausryhmään ovat osallistuneet Nina Karasmaa YTV:ltä ja Pekka Räty Tiehallinnosta.

Haluan kiittää Virpi Pastista avusta koko diplomityöprojektin edistämisessä, aiheiden keksimisessä ja rahoituksen hankkimisessa sekä lopulta työn ohjaamisesta. Hannu Lehdolle kiitokset etenkin henkisestä tuesta työn aikana; on tärkeää tietää, että voi kysyä mitä vain mistä vain, jos eteen tulee ongelmia.

Ideoista ja palautteesta työn aikana kiitän professori Ernvallia, Nina Karasmaata ja Pekka Rätystä; kokoukset veivät työtä hyppäyksinä eteenpäin. Lisäksi haluan erityisesti kiittää kaikkia kenttätutkimukseen osallistuneita sukulaisiani, opiskelukavereitani, työtovereitani ja muita vapaaehtoisia: tämä työ ei olisi ollut mahdollinen ilman teitä. Lopuksi vielä kiitokset tuesta perheelleni ja työtovereilleni koko diplomityöprosessin ajalta.

Työn päättyessä vuoden 2009 lopussa jätetään jäähyväiset diplomityöni ja opiskelijaelämäni lisäksi monelle organisaatiolle. Liikennelaboratorio lakkasi olemasta jo jonkin aikaa sitten, ja vuoden vaihtuessa lakkaavat nykymuodossa olemasta niin Teknillinen korkeakoulu, YTV kuin Tiehallintokin.

Espoossa 30.11.2009

Annika Rantala

Sisällysluettelo

	Sivu
TIIVISTELMÄ	2
ABSTRACT	3
ALKUSANAT	4
SISÄLLYSLUETTELO.....	5
TAULUKKOLUETTELO	8
LIITELUETTELO	9
TERMIT JA KÄSITTEET	10
1 JOHDANTO	13
2 KÄYTETYT MENETELMÄT JA TYÖN RAJAUS	14
3 KIRJALLISUUSSELVITYS	16
3.1 Satelliittipaikannus ja sen käyttö liikkumistutkimuksissa	16
3.1.1 GPS-tekniikan perusteet.....	16
3.1.2 GPS-laitteet liikkumistutkimuksissa	20
3.2 GPS:n käyttö liikkumistutkimuksissa	26
3.2.1 Paikannus ja muut uudet tekniikat liikkumisen tutkimisessa.....	26
3.2.2 Katsaus toteutettuihin GPS-avusteisiin liikkumistutkimuksiin.....	34
3.3 Algoritmit GPS-aineistolle.....	43
3.3.1 GPS-aineiston automaattinen käsittely	43
3.3.2 Kulutavanpäättelyalgoritmit.....	49
4 TUTKIMUKSEN TOTEUTUS	54
4.1 Kenttätutkimus	54
4.1.1 Laitteiden valinta ja hankkiminen	54
4.1.2 Kenttätutkimuksen suunnittelu	56
4.1.3 Kenttätutkimuksen suorittaminen	59
4.2 GPS-aineiston käsittely ja analysointi.....	60
4.3 Algoritminkehitys	63
4.3.1 Algoritmien valinta	63
4.3.2 Algoritmien toteutus ja testaus.....	66
5 TULOKSET	72
5.1 Kenttätutkimus	72
5.2 Algoritminkehitys	77
6 POHDINTAA	82
7 JOHTOPÄÄTÖKSET.....	85
7.1 Kirjallisuusselvitys.....	85
7.2 Kenttätutkimus	86
7.3 Algoritminkehitys	87
7.4 Kehityspolkuja	87
8 YHTEENVETO	90
LÄHDELUETTELO.....	93
LIITTEET	96

KUVALUETTELO

Kuva 1. GPS-järjestelmän osat: 1) kontrolliverkko (eli valvonta-asemat, <i>monitoring stations</i>), 2) satelliitit (<i>space segment</i>) ja 3) käyttäjät (<i>user segment</i>).	17
Kuva 2. Korkeat rakennukset ja puut häiritsevät usein GPS-signaalin vastaanottamista.	19
Kuva 3. Yksi ensimmäisistä liikkumistutkimuksia varten kehitetyistä kannettavista GPS-laitteista, jota käytettiin muun muassa vuonna 2002. Vieressä matkapuhelin koon havainnollistamiseksi. (Lähde: Stopher ym. 2008.).....	22
Kuva 4. Seuraavan sukupolven GPS-laite vuodelta 2005. Vieressä matkapuhelin koon havainnollistamiseksi. (Lähde: Stopher ym. 2008.).....	22
Kuva 5. Tutkimuksessa käytetty GPS-laite (iBT747A+) vuodelta 2008. Laitteen ulkomitat ovat 72,2x46,5x20 mm ja paino noin 70 g.	22
Kuva 6. Esimerkki tutkimuksessa käytetyn GPS-laitteen (iBT747A+) tallentamista tiedoista. (Vasemmalta oikealle: tietueen järjestysnumero, päivämäärä (UTC), UTC-kellonaika, leveysaste, pallonpuolisko (pohjoinen N / eteläinen S), pituusaste, pallonpuolisko (itäinen E / läntinen W), korkeusasema ja nopeus.).....	24
Kuva 7. GPS:llä kerättävät tiedot. Kuvassa näkyy kuljettu reitti (ja joidenkin pisteiden koordinaatit), ja tietyssä pisteessä tiedoston nimi, päivämäärä, kellonaika, matkan alusta kulunut aika, nopeus ja korkeusasema.	28
Kuva 8. GPS-laitteen tallentama ”reitti” kun laite on ollut useita tunteja sisätiloissa paikoillaan valkoisen nuolen osoittamassa kohdassa. Mustalla on merkitty rakennuksen ääriiviivat. Punaiset pallot kuvaavat kohtia, joissa laitteessa olevaa nappulaa on painettu matkan alun ja lopun merkiksi. (Matkan alkua kuvaavat pallot eivät välttämättä ole kovin tarkasti oikeassa kohdassa, koska laite on juuri tuotu ulos rakennuksesta ja paikannustarkkuus on siis siksak-kuvion luokkaa.) Siksak-kuvion koko vaihtelee eri rakennuksissa sen mukaan, miten hyvin GPS-signaali kuuluu rakennuksen sisään.	45
Kuva 9. Erot GPS-laitteen määrittämän nopeuden (siniset pisteet) ja koordinaattien muutoksesta lasketun nopeuden (punaiset pisteet) välillä eräällä bussimatalla.	61
Kuva 10. Esimerkki nopeus- ja kiihtyvyysjakaumasta erään kenttätutkimukseen osallistuneen keräämässä aineistossa. Schuessler ja Axhausen käyttivät analyysissään samanlaisia kuvia. Tämä jakauma on tässä työssä kerätystä aineistosta.	62
Kuva 11. Sama nopeus- ja kiihtyvyysjakauma kuin edellisessä kuvassa, mutta nyt on mukana tieto siitä, mitä kulkutapaa mikin piste edustaa. Nopeusalueella 80–120 km on junaa merkitsevien pisteiden alla piilossa vähintään yhtä paljon autoa merkitseviä pisteitä.	62
Kuva 12. Eri kulkutavoilla tehdyt matkan osat aineistossa.	72
Kuva 13. Osallistujien sukupuoli	73
Kuva 14. Osallistujien ikä.....	73
Kuva 15. Osallistujien kotikunta.....	73
Kuva 16. Osallistujien työ-/opiskelupaikan kunta	74
Kuva 17. Kokemukset tutkimuksen kestosta	74
Kuva 18. Halukkuus osallistua uuteen tutkimukseen	75
Kuva 19. Tutkimuksen rasittavuus.....	75

Kuva 20. Tutkimukset aikana esiintyneet vaikeudet.....	75
Kuva 21. Matkan osien määrät eri kulkutavoilla aineistossa (sininen) sekä kahden algoritmin löytämät kulkutavat. Vasemmanpuoleinen sininen pylväs kuvaa kävelymatkojen määrää aineistossa. Heti sen oikealla puolella oleva puna-harmaa pylväs kuvaa päättelysääntöalgoritmin (versio 1.11) löytämiä kävelymatkoja. Pylvään punainen osuus kuvaa oikein tunnistettuja kävelymatkoja ja harmaa osuus muita kuin kävelymatkoja, jotka algoritmi on kuitenkin tulkinnut kävelymatkoiksi. Vihreä-harmaa pylväs kuvaa vastaavasti pisteytysalgoritmin (versio 2.0) löytämiä kävelymatkoja. Samoin tulkitaan muut kulkutavat.	78
Kuva 22. Eri kulkutavoilla tehtyjen matkan osien määrät aineistossa ja algoritmeilla. Mukana myös väärin tunnistetut matkan osat, jotta saadaan käsitys siitä, mikä on todellinen tulos algoritmia käytettäessä. 1, 5 ja 10 s selitteessä tarkoittavat, että 1 s:n kohdalla on käytetty koko GPS-aineistoa, 5 s:n kohdalla joka viidettä aineiston tietuetta ja 10 s:n kohdalla joka kymmenettä tietuetta.	79
Kuva 23. Eri kulkutapojen osuudet matkan osilla aineistossa ja algoritmeilla. Mukana myös väärin tunnistetut matkan osat, jotta saadaan käsitys siitä, mikä on todellinen tulos algoritmia käytettäessä. 1, 5 ja 10 s selitteessä tarkoittavat, että 1 s:n kohdalla on käytetty koko GPS-aineistoa, 5 s:n kohdalla joka viidettä aineiston tietuetta ja 10 s:n kohdalla joka kymmenettä tietuetta.	80
Kuva 24. Matkan osien määrät eri kulkutavoilla validointiaineistossa (sininen), sekä kahden algoritmin (päättelysääntöversio 1.11 ja pisteytysversio 2.0) löytämät kulkutavat. Tulkintaohje on sama kuin kuvassa 21. Tämä kuva perustuu siis riippumattomaan lisäaineistoon, jota ei käytetty algoritmin tekemiseen.	81
Kuva 25. Kaavakuva joukkoliikennevälineiden tunnistamisesta reittien perusteella. Oranssit pallot kuvaavat asemia tai pysäkkejä ja katkoviivat niiden välissä summittaisia reittejä. Punainen viiva kuvaa alueen, jolla kulkuvälineen käyttö ylipäättään on mahdollista (olettaen siis että kuvassa ovat kyseisen välineen kaikki pysäkit). Sinisen viivan sisään jää käytävä, johon GPS-aineiston koordinaatteja voidaan verrata, jos käytettävissä on reitin koordinaatit.	88

TAULUKKOLUETTELO

Taulukko 1. GPS-laitteen käynnistystavat	20
Taulukko 2. Yhteenvedo tiedoista, joita GPS-laitteen avulla voidaan ja ei voida kerätä. (Lähde: Bonsall ym. (2006), Table 1.) Taulukon loppuun on tehty joitain huomautuksia (*-***), joita ei ole alkuperäisessä lähteessä.	30
Taulukko 3. Muutamien GPS-tutkimusten ominaisuuksia esimerkinomaisesti esiteltynä	35
Taulukko 4. Tekijöitä, jotka pienentävät osallistumistodennäköisyyttä GPS-tutkimukseen. Eri tutkimuksissa tekijöitä on vertailtu eri tavoin, eivätkä ne ole yhteismitallisia, mutta taulukko antaa osviittaa siitä, millaisia henkilökohtaisia ominaisuuksia ja liikkuvuustietoja kannattaa ottaa huomioon tutkimuksia suunniteltaessa ja analysoitaessa.....	40
Taulukko 5. Vastaajien kokema tutkimuksen eri osien kuormittavuus hollantilaisessa GPS-tutkimuksessa. Kuormitusta arvioitiin kolmesta eri asiasta: GPS-laitteen kuljettaminen mukana, GPS-laitteen lataaminen ja internetsovelluksen käyttäminen. Vastausvaihtoehdot olivat hyvin kuormittavaa, jossain määrin kuormittavaa ja ei kuormittavaa. (Lähde: Bohte ja Maat 2008, Table 3.)	42
Taulukko 6. Chungin ja Shalabyn (2005) kulkutavanpäättelyalgoritmin taustatiedot....	49
Taulukko 7. Bohten ja Maatin (2008) kulkutavanpäättelyalgoritmin taustatiedot	50
Taulukko 8. Schuesslerin ja Axhausenin kulkutavanpäättelyalgoritmin taustatiedot.....	51
Taulukko 9. Stopherin ym. (2008) kulkutavanpäättelyalgoritmin taustatiedot	52
Taulukko 10. Kulkutavanpäättelyalgoritmin eri versioissa käytetyt muuttujat	67
Taulukko 11. Tulokset kulkutavanpäättelyalgoritmin versiosta 0.0. Mahdollisimman yksinkertainen, keskinopeuksiin perustuva algoritmi.	68
Taulukko 12. Tulokset kulkutavanpäättelyalgoritmin versiosta 1.0. Lisätietona joukkoliikenteen pysäkkien paikat.	68
Taulukko 13. Tulokset kulkutavanpäättelyalgoritmin versiosta 1.11. Lisätietona joukkoliikenteen pysäkkien paikat, nopeuden 90. persentiili ja nopeuden keskihajonta.	69
Taulukko 14. Tulokset kulkutavanpäättelyalgoritmin versiosta 2.0. Käytetty pisteytysjärjestelmää, jossa muuttujina keskinopeus, nopeuden 90. persentiili ja nopeuden keskihajonta, joukkoliikenteen pysäkit ja matkan aikana tehdyt pysähdykset.	70
Taulukko 15. Vertailu päättelysääntöjen ja pisteytysjärjestelmän algoritmien toimivuudesta.	78

LIITELUETTELO

Liite 1. GPS-laitteiden vertailu	96
Liite 2. Tiedonsiirtoyhteyksien hintoja (8.10.2009).....	100
Liite 3. Tutkimuslomakkeet	101
Liite 4. Kenttätutkimuksen osallistujille jaettu ohjeistus (3 sivua) tutkimuksen suorittamisesta ja GPS-laitteen käyttämisestä.....	109
Liite 5. Nopeus- ja kiihtyvyysskuvaajia, joissa on esitetty sekä suoraan GPS-laitteesta saatu nopeus että koordinaattien muutoksen perusteella laskettu nopeus, tai vastaavasti näiden pohjalta lasketut kiihtyvyydet. (Koordinaattien perusteella laskettu nopeus punaisella, GPS-laitteesta saatu nopeus sinisellä.).....	112
Liite 6. Versioita päättelysääntöihin perustuvasta kulkutavanpäättelyalgoritmista.....	116
Liite 7. Kulkutavanpäättelyalgoritmin pisteytysjärjestelmä ja sen perustelut	118
Liite 8. Aineistonjakamisalgoritmin luonnos, vielä kokeiluvaiheessa.....	122
Liite 9. Tunnistustuloksia pisteytysjärjestelmästä	123
Liite 10. Tulokset testistä, jossa poimittiin GPS-aineistosta joka 5. tai joka 10. tietue	125
Liite 11. Tulokset validointiaineistolla tehdyistä testeistä	127

TERMIT JA KÄSITTEET

Selityksissä esiintyvät alleviivatut sanat löytyvät listasta omana hakusananaan.

Aktiviteetti

Aktiviteetilla tarkoitetaan tässä työssä mitä tahansa toimintaa tai oleskelua, joka ei ole siirtymistä paikasta toiseen. Tutkimuksen näkökulmasta ihmisen päivä siis jakautuu aktiviteetteihin ja matkoihin, joiden tarkoituksena on siirtyä paikasta toiseen suorittamaan seuraavaa aktiviteettiä. Kärjistäen päivä voisi jakaantua esimerkiksi seuraavasti: aktiviteetti: aamutoimet – matka: bussilla työpaikalle – aktiviteetti: työ – matka: bussilla kuntosalille – aktiviteetti: kuntosalilla käynti – matka: kävellen kotiin – aktiviteetti – ilta kotona ja nukkuminen.

Algoritmi

Äärellinen joukko täsmällisiä ohjeita, joita seuraamalla voidaan ratkaista tietty ongelma.

Geokoodaus

Ks. koordinaattihaku

GPS (*Global Positioning System*)

USA:n omistama koko maapallon kattava satelliittipaikannusjärjestelmä.

GPS-aineisto

Tässä työssä GPS-loggerin tallentama raakadata, käytännössä jono pituus- ja leveyskoordinaatteja. Koordinaatteihin liittyy yleensä myös lisätietoja kuten päivämäärä, kellonaika, korkeusasema ja hetkellinen nopeus (ks. kuva 6).

GPS-laite

Tässä tutkimuksessa laite, joka on sekä GPS-vastaanotin että GPS-loggeri, toisin sanoen sekä paikantaa itsensä että tallentaa paikannetun sijainnin.

GPS-loggeri (puhekieltä; engl. *GPS data logger*, *GPS data logging device*)

Laite, johon tallentuu kuljettu reitti (koordinaatit) vastaanotetun GPS-signaalin perusteella. Usein sama laite sekä paikantaa sijainnin että tallentaa kuljetun reitin, mutta GPS-signaali voi tulla myös erillisestä vastaanottimesta.

GPS-vastaanotin

Laite joka sisältää GPS-antennin (tai puhekielellä GPS-sirun) ja pystyy siten vastaanottamaan GPS-satelliittien signaalia ja päättelemään siitä sijaintinsa maapallolla.

Henkilöliikennetutkimus

Liikkumistutkimus, jossa yleensä pyritään saamaan kattava kuva tietyn alueen (esim. valtio, kaupunkiseutu) asukkaiden liikkumisesta, kuten käytetyistä kulkutavoista ja matkojen määristä, pituuksista, kestoista ja tarkoituksista. Tuloksia käytetään usein muun muassa tilastointiin ja liikennemalleihin.

Koordinaattihaku (puhek. usein geokoodaus, engl. *geocoding*)

Useimmiten koordinaattihaku viittaa siihen, että osoitteelle etsitään sitä mahdollisimman hyvin vastaavat koordinaatit. Yleisemmin voidaan muuhunkin maantieteelliseen tietoon kuin osoitteeseen (esimerkiksi postinumeroon) liittää maantieteellinen tunniste, kuten koordinaatit, lähin risteys, esimerkiksi liikennesuunnittelussa käytettävä osa-alue tai vastaava (esimerkiksi Greaves 2004).

Koordinoitu yleisaika, UTC-aika (engl. *UTC time, Coordinated Universal Time*)

”Koordinoitu yleisaika, *Coordinated Universal Time*. Normaalisti käytettävä aika, joka seuraa kansainvälistä atomiaikaa, mutta kun ero Greenwichin keskiaurinkoaikaan nousee yli 0,9:ksi sekunniksi, UTC-aikaa siirretään tasan yhden sekunnin eteen- tai taaksepäin *karkaussekunnilla*”¹. Tässä työssä käytetyt GPS-laitteet tallentavat UTC-kellonajan. Käytännössä UTC-aika on liikennesuunnittelun näkökulmasta sama kuin Greenwichin keskiaurinkoaika GMT, eli aurinkoaika nollameridiaanilla. GPS-laitteita käytettäessä on kuitenkin muistettava, että UTC-aika ei siirry kesäaikaan, eli Suomen aika on talvella 2 tuntia UTC-aikaa edellä ja kesäajan voimassa ollessa 3 tuntia UTC-aikaa edellä, jolloin GPS-aineistoa on korjattava asianmukaisesti, kun UTC-aikaa muutetaan paikalliseksi ajaksi.

Liikkumistutkimus

Tässä työssä liikkumistutkimusta käytetään yleiskäsitteenä kaikenlaisille ihmisten liikkumista, liikkumistottumuksia ja niiden muutoksia selvittäville tutkimuksille; enimmäkseen asioita tarkastellaan kuitenkin henkilöliikennetutkimusten näkökulmasta.

Matka

”Matka on siirtymistä paikasta toiseen, esimerkiksi kotoa kauppaan tai työpaikalle. Meno ja paluu ovat erillisiä matkoja. Matkoiksi luetaan kaikki matkat, myös lyhyet, jos ne ulottuvat pihapiirin ulkopuolelle. Pihapiiri voi tarkoittaa vastaajan oman kodin pihapiiriä tai muuta sen hetkistä oleskelupaikkaa. Matkoiksi ei lasketa liikkumista omassa pihapiirissä tai tilalla eikä ammattiautoilijoiden ja muiden liikennevälineissä työskentelevien työssään tekemiä matkoja.” (Henkilöliikennetutkimus 2004–2005)

Opt-in (”päättää tulla mukaan”)

GPS-tutkimuksessa opt-in-vaihtoehtoa käytetään kuvaamaan sitä, että ihmisiltä kysytään, haluavatko he osallistua GPS-tutkimukseen. Ne, jotka haluavat osallistua, ilmoittavat päättäneensä että haluavat mukaan. (Markkinoinnissa opt-in tarkoittaa, että markkinointi on sallittua vain etukäteen annetun luvan kanssa.)

Ks. myös vastakohta opt-out.

Opt-out (”jättäytyä pois”)

GPS-tutkimuksessa opt-out-vaihtoehto kuvaa sitä, että ihmisille kerrotaan, että heidät on valittu GPS-tutkimukseen. Ne, jotka eivät halua osallistua, voivat kieltäytyä eli jättäytyä pois tutkimuksesta. Perinteiset matkapäiväkirjatutkimukset, joissa esimerkiksi koko maasta arvotaan otos, jolle lähetetään tutkimuslomakkeet etukäteen ilmoittamatta, ovat opt-out-tutkimuksia. (Markkinoinnissa opt-out tarkoittaa, että markkinointi on sallittua, ellei sitä ole erikseen kielletty.)

Ks. myös vastakohta opt-in.

Osallistujiin kohdistuva kuormitus (vapaa käännös engl. termistä *respondent burden*)

Rasitus ja vaiva, jota tutkimukseen osallistuvilla aiheutuu tutkimukseen osallistumisesta ja siihen liittyvien tehtävien suorittamisesta. Kuormitusta voidaan mitata sillä, kuinka moni suostuu osallistumaan tutkimukseen ja sillä, kuinka hyvin tutkimukseen liittyvät tehtävät suoritetaan (Bricka 2008).

¹ Lähde Wikipedia <http://fi.wikipedia.org/wiki/Aikaj%C3%A4rjestelm%C3%A4#UTC>

Paikkatieto (engl. *geographic information*)

Paikkatieto on tietoa kohteista, joiden paikka Maan suhteen tunnetaan². Toisaalta paikkatieto on paikannettua kohdetta tai ilmiötä kuvaava sijaintitiedon ja ominaisuustiedon looginen tietokokonaisuus³. Tässä työssä oleellisia paikkatietoja ovat esimerkiksi tiedot joukkoliikenteen pysäkkien sijainnista.

Persentiili

”Persentiili eli sadannes- tai prosenttipiste kuuluu ns. fraktiileihin eli jakauman osuuspisteisiin. Se ilmoittaa muuttujan arvon, jonka alapuolelle jakaumassa jää tapauksista 1 % (1. persentiili), 2 % (2. persentiili), 15 % (15. persentiili) jne.”⁴

Prompted recall (*valitettavasti ei suomenkielistä termiä*)

Menetelmä jossa osallistujan matkoista kerätyt tiedot käydään myöhemmin läpi uudelleen ja kysellään tarkennuksia ja korjauksia; tiedot voidaan näyttää esim. kartoilla ja taulukoina, jotka auttavat osallistujaa muistamaan, missä hän on käynyt ja mitä tehnyt.

Satelliittipaikannus

Sijainnin maapallolla (koordinaattien) päättelemisen satelliittijärjestelmästä saadun tiedon avulla.

Sumea logiikka (engl. *fuzzy logic*)

”Sumea logiikka on matemaattisen logiikan laajennus, jossa propositiolla on diskreetin totuusarvon (tosi tai epätosi) sijasta reaalinen totuusarvo suljetulla välillä nollasta yhteen” (http://fi.wikipedia.org/wiki/Sumea_logiikka). Lähtökohdat ymmärtämiselle ja listan lähdekirjallisuutta tarjoaa esimerkiksi englanninkielinen wikipedia⁵. Tässä työssä ei ole käytetty varsinaista sumea logiikkaa, vaan yksinkertaisempaa pisteytysjärjestelmää, joten sumean logiikan perusteita ei käydä sen tarkemmin läpi.

² Lähde Geoinformatiikan sanasto, termi 37, Sanastokeskus TSK, Helsinki 2005, online <http://www.tsk.fi/fi/info/GeoinformatiikanSanasto.pdf>

³ Lähde Paikkatietotekniikan sanasto 2004

⁴ Lähde: Tilastokeskus: <http://www.stat.fi/meta/kas/persentiili.html>

⁵ http://en.wikipedia.org/wiki/Fuzzy_logic

1 JOHDANTO

Ihmisten liikkumistottumuksia on eri maissa jo pitkään kartoitettu erilaisilla tutkimuksilla. Tutkimuksilla selvitetään liikenteen ja liikkumisen nykytilaa, eli muun muassa mitä kulkutapoja käytetään, mistä mihin matkoja tehdään ja mihin aikaan liikutaan.

Liikkumistutkimusten tuloksista kootaan tilastoja ihmisten liikkumisesta eri käyttötarkoituksiin. Lisäksi tutkimusten tuloksia käytetään muun muassa liikenteen mallintamiseen, liikenne-ennusteiden tekemiseen, liikennesuunnitteluun ja päästömallien laadintaan. Ruuhkautuminen, saasteet ja ilmastonmuutos ovat muutamia ongelmia, joiden ratkaisemisessa liikkumistottumusten, niiden muutosten ja tottumuksiin vaikuttavien tekijöiden tuntemisen toivotaan auttavan.

Liikkumisen tutkimisen menetelminä ovat pitkään olleet paperiset matkapäiväkirjat ja haastattelut. Niiden ongelmana on kuitenkin, että ihmiset usein unohtavat ilmoittaa osan matkoistaan, ja esimerkiksi tiedot matkan pituudesta ja matka-ajasta arvioidaan yleensä väärin. Lisäksi tutkimuksiin osallistuminen on vaivalloista vastaajille, joten tutkimuksia, joissa samaa ihmistä seurattaisiin pidemmän aikaa, on hankala toteuttaa. Perinteisten tutkimusten vastausprosentit ovat myös pienentyneet ajan saatossa.

Viimeisten 15–20 vuoden aikana on vähitellen kehitetty uusia menetelmiä ja elektronisia apuvälineitä. Ne voivat olla esimerkiksi laitteita, jotka mittaavat matkan pituutta ja kestoa tai laitteita, joihin vastaaja jo matkan alkaessa merkitsee matkan ominaisuuksia, kuten matkan tarkoituksen, kulkutavan ja mukana olevien henkilöiden lukumäärän. Ensimmäiset apuvälineet ovat usein olleet ajoneuvoon, käytännössä omaan autoon sidottuja, koska silloin matkan pituus ja nopeus saadaan automaattisesti, ja laite saa virtaa autosta.

Yksi uusista menetelmistä on paikannus, joka GPS-sovellusten yleistymisen myötä on noussut esiin myös liikkumistutkimusten apukeinona. Paikannustekniikan kehittyminen on tehnyt erilaiset kokeilut kannattaviksi; paikannuslaitteet ovat entistä halvempia ja pienempiä ja saatava tieto on entistä tarkempaa. Parhaassa tapauksessa paikannuksella voidaan saada ihmisten liikkumisesta paljon tietoa ilman, että tutkittavien tarvitsee tehdä muuta kuin kantaa laitetta mukanaan. Tähän tarvitaan kuitenkin pitkälle kehitettyjä laitteita ja ohjelmistoja, joilla tietoa kerätään, muokataan ja tulkitaan.

Tämän diplomityön tarkoitus on olla yksi osa tuota paikannuksen hyödyntämiseen liittyvää kehitystyötä. Työ koostuu kolmesta vaiheesta: Ensin tehdään katsaus ulkomaisiin tutkimuksiin, joissa GPS-paikannusta on hyödynnetty liikkumistutkimuksissa. Erityisesti keskitytään tapauksiin, joissa paikannusaineistosta on jälkikäteen pyritty tunnistamaan käytetty kulkutapa. Toinen osa on pienimuotoinen pilottitutkimus, jossa kerätään GPS-aineistoa pääkaupunkiseudun asukkaiden matkoista. Kerättyä aineistoa käytetään kolmannessa vaiheessa kehitettäessä algoritmia, jolla kulkutapa voidaan Suomen oloissa tunnistaa GPS-aineistosta. Tavoitteena on ohjelmoida algoritmi toimivaksi ohjelmaksi ja testata sen tarkkuutta.

Kulkutavan päättelemisen ei tietenkään yksinään riitä kovin pitkälle, mutta se on tässä työssä valittu ensimmäiseksi askeleeksi kohti tietojen automaattista tulkintaa. Seuraavissa vaiheissa – esimerkiksi reittien ja matkan tarkoituksen päättelyssä – tarvitaan jo laajempaa paikkatieto-osaamista.

2 KÄYTETYT MENETELMÄT JA TYÖN RAJAUS

Työssä käytettiin menetelminä kirjallisuusselvitystä, kenttätutkimusta ja algoritmin kehittämistä. Seuraavassa on lyhyesti kuvattu työn eri vaiheet. Samalla on käsitelty työn rajausta.

Kirjallisuusselvitys

Kirjallisuusselvityksen tarkoituksena oli perehtyä lähinnä artikkelien ja kansainvälisissä konferensseissa esiteltyjen tulosten avulla GPS-paikannukseen, sen käyttämiseen liikkumistutkimuksissa sekä jo kehitettyihin algoritmeihin, joilla GPS-aineistosta voidaan päätellä muun muassa kulkutapa. Suurin osa käytetystä kirjallisuudesta ja muusta aineistosta on englanninkielistä; muun kuin suomen- ja englanninkielistä aineistoa ei aktiivisesti etsitty.

Varsinaista paikannuksen tekniikkaa käsitellään kirjallisuussosiossa hyvin lyhyesti, sillä paikannus on hyvin laaja tieteenala, josta satelliittipaikannus ja sen osana GPS-paikannus muodostavat vain yhden osan. Paikannustekniikan voidaan katsoa kuuluvan enemmän maanmittaukseen kuin liikennetekniikkaan, joten tässä keskitytään paikannustekniikan sovelluksiin liikennetekniikassa. GPS-paikannuksen lisäksi sivutaan hieman muita paikannustekniikoita, joita voidaan soveltaa liikenne- ja liikkumistutkimuksissa.

GPS-tekniikan lisäksi toisen laajemman tarkastelukontekstin aiheelle tarjoavat liikkumistutkimukset. Kirjallisuussosiossa käsitellään hieman liikkumistutkimusten tavoitteita ja nykytilaa taustaksi sille, miksi niiden avuksi on ryhdytty kehittämään erilaisia teknisiä apuvälineitä. Lähinnä tarkastellaan henkilöliikennetutkimuksia, koska ne ovat usein laajoja ja toistuvia, jolloin menetelmien kehittäminen on kannattavaa. Myös joitain muita aputekniikoita kuin GPS-paikannus mainitaan lyhyesti, mutta niihin ei perehdytä tarkemmin.

GPS-aineiston keräämisestä ja hyödyntämisestä saa hyvän kuvan tutustumalla toteutettuihin tutkimuksiin. Koska kirjallisuusselvitys toimi myös pohjana kenttätutkimuksen suunnittelussa, kirjallisuussosiossa käsitellään jonkin verran myös toteutettujen kenttätutkimusten yksityiskohtia sekä GPS-laitteita.

GPS-aineiston prosessointi on monista osista koostuva kokonaisuus, johon tarvitaan erilaisia algoritmeja. Tässä työssä keskityttiin kulkutavanpäättelyalgoritmeihin. Kirjallisuusselvityksessä käsitellään lyhyesti myös muita algoritmeja, jotka tukevat tai täydentävät kulkutavanpäättelyä.

Kenttätutkimus

Kenttätutkimuksella oli kaksi tarkoitusta: tuottaa aineistoa kulkutavantunnistus-algoritmin kehittämiseen ja testaamiseen sekä toimia pilottitutkimuksena Suomessa tulevaisuudessa GPS-laitteiden avulla toteutettavia liikkumistutkia ajatellen.

Työn suppeuden takia pilotti oli varsin rajattu. Osallistujia oli 50 ja he olivat pääosin pääkaupunkiseudulta. Siten myös suurin osa matkoista tehtiin pääkaupunkiseudun sisällä.

Pilotissa vapaaehtoiset kantoivat noin 2 päivän ajan mukanaan laitetta, joka vastaanotti GPS-signaalia ja tallensi kuljetun reitin. Samalla osallistujat täyttivät paperista matkapäiväkirjaa, jossa oli GPS-aineiston tulkintaa helpottavia kysymyksiä. Tavoitteena oli, että vapaaehtoiset käyttäisivät kattavasti eri kulkutapoja, mutta varsinaisen kenttätutkimuksen päätyttyä kerättiin vielä lisähavaintoja niistä kulkutavoista, joilla tehtyjä matkoja aineistossa oli vähän (raitiovaunu ja metro).

Tutkimuspäivien jälkeen GPS-laitteiden tallentamat tiedot purettiin tietokoneelle ja paperilomakkeilla kerätyt tiedot koodattiin sähköiseen muotoon, ja ne muokattiin sopivaan muotoon jatkohyödyntämistä varten.

Algoritminkehitys

Kulutusalan päättelyalgoritmin kehityksessä pyrittiin kirjallisuusselvityksessä läpikäytyjen algoritmien ideoita hyödyntämällä ja yhdistelemällä kehittämään algoritmi, joka tunnistaisi GPS-aineistosta eri kulkutavat. Kenttätutkimuksesta saatua aineistoa käytettiin eri kulkumuotojen ominaisuuksien tarkasteluun ja vertailuun.

Lähtökohtana työssä oli algoritmin kehittäminen ainoastaan kulkutapojen tunnistamista varten. Työn ulkopuolelle rajattiin esimerkiksi kirjallisuusosassa mainitut matkan tarkoituksen päättelyyn kykenevät algoritmit tai algoritmit, jotka osaavat sijoittaa kuljetun reitin tieverkolle, koska ne olisivat vaatineet paljon lisää työtä ja paikka-tietoaineistojen ja -ohjelmien käyttöä.

Lähtötietona käytettiin joukkoliikenteen pysäkkien sijaintitietoja. Myös muita paikkatietoaineistoja, kuten joukkoliikenteen reittejä, tieverkkoja ja maankäyttötietoja on saatavilla, mutta ne rajattiin työn ulkopuolelle.

Edellytyksenä kulutusalan tunnistamiselle on, että GPS-laitteen tallentama koordinaattijono jaetaan ensin osiin, joissa on käytetty vain yhtä kulkutapaa. Siksi kehiteltiin resurssien puitteissa alustavaa algoritmia myös tähän tarkoitukseen.

Algoritmien toimivuuden testaamista varten ne toteutettiin C-kielisiksi ohjelmiksi.

3 KIRJALLISUUSSELVITYS

3.1 Satelliittipaikannus ja sen käyttö liikkumistutkimuksissa

3.1.1 GPS-tekniikan perusteet

Satelliittipaikannusjärjestelmät

Satelliittipaikannusta käydään tässä läpi liikkumistutkimusten näkökulmasta siltä osin, miten tarkkaa tietoa sen avulla voidaan saada ja millaisia virheitä aineistoon voi tulla. Tarkempia teknisiä yksityiskohtia ei ole tässä enempää kuvailtu, koska satelliittipaikannus itsessään on varsin laaja tieteenala, ja yksityiskohtia voi tarvittaessa selvittää muista lähteistä (perusteet satelliittipaikannuksesta esimerkiksi Leick 2004, GPS-tekniikan mahdollisuudet liikkumis- ja liikennetutkimusten kannalta esimerkiksi Wolf 2004).

Satelliittipaikannukseksi voidaan kutsua mitä tahansa satelliittien avulla toteutettua paikannusjärjestelmää. GPS (*Global Positioning System*) on vain yksi maailmanlaajuisista satelliittipaikannusjärjestelmistä. GPS on yhdysvaltalainen, alun perin sotilaskäyttöön kehitetty järjestelmä. Muita maailmanlaajuisia tai sellaisiksi tarkoitettuja satelliittipaikannusjärjestelmiä ovat venäläinen GLONASS sekä kehitteillä oleva EU:n Galileo. (Bonsall ym. 2006.)

Galileo on kärsinyt rahoitusvaikeuksista, ja se saataneen täyteen toimintavalmiuteen aikaisintaan vuonna 2013 (Gibbons 2008b). Kiina kehittää lisäksi omaa satelliittipaikannusjärjestelmäänsä, joka tunnetaan nimellä Beidou tai Compass (Gibbons 2008a). Tällä hetkellä GPS on käytännössä ainoa toimiva vaihtoehto. Siksi tässä työssä samoin kuin suurimmassa osassa läpikäydystä kirjallisuudesta keskitytään vain GPS:ään ja unohdetaan muut järjestelmät kunnes ne saadaan toimintaan. On kuitenkin huomattava, että suunniteltaessa pitkäkestoisia projekteja – kuten henkilöliikennetutkimukset – Euroopassa, myös Galileon olemassaolo kannattaa huomioida (Bonsall ym. 2006).

GPS-tekniikka

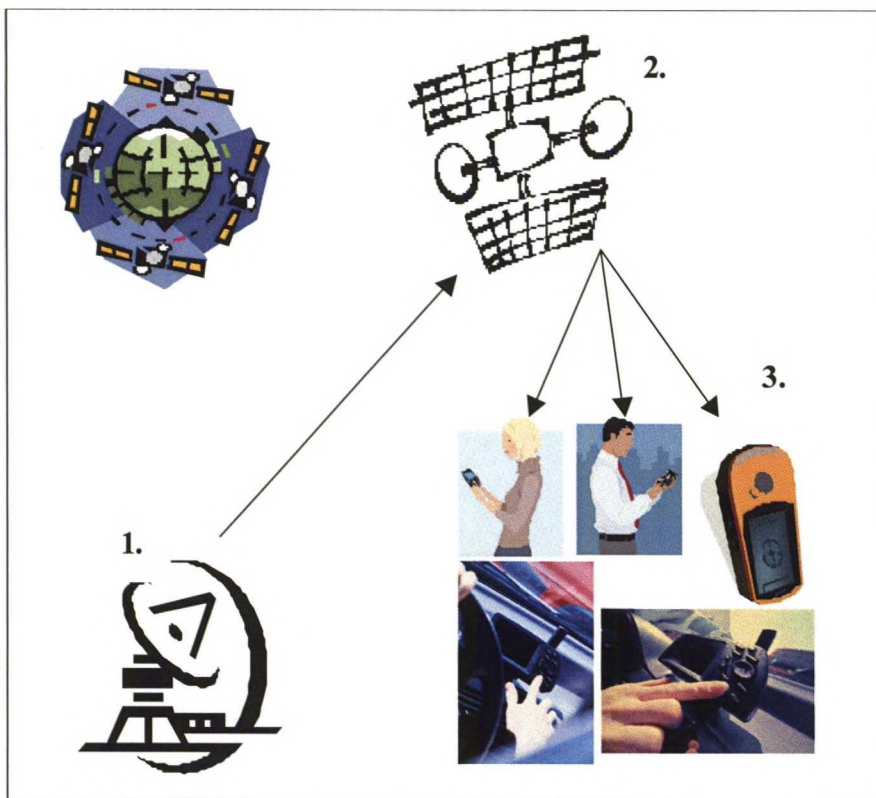
GPS-järjestelmä koostuu kolmesta osasta: satelliiteista, kontrolliverkosta ja käyttäjistä (kuva 1). Satelliitteja tarvitaan toimivaan järjestelmään vähintään 24, ja ne kiertävät maapalloa tarkasti laskettuja ratoja pitkin. Satelliiteissa on hyvin tarkat atomikellot, ja ne lähettävät radioaalloilla signaalia käyttäjien GPS-vastaanottimiin. Kontrolliverkko koostuu valvonta-asemista, joilta tarkkaillaan satelliittien tilaa ja tehdään tarvittavat muutokset ja korjaukset. (Leick 2004.)

GPS-satelliittien lähettämässä signaalissa on atomikellon aika ja navigaatiosignaali. GPS-laite vastaanottaa signaalia useammalta satelliitilta yhtä aikaa. Navigaatiosignaali sisältää muun muassa tarkan tiedon satelliitin lentoradasta (efemeris) eli satelliitin sijainnin ajan funktiona, tiedon satelliitin tilasta (onko satelliitti kunnossa eli lähettääkö se oikeaa tietoa) sekä almanakan, jossa on muun muassa karkeasti kaikkien järjestelmän satelliittien lentoratatiedot ja tilatiedot. (Leick 2004.)

Yksinkertaisimmillaan paikannus tapahtuu siten, että vastaanotetun ajan perusteella GPS-vastaanotin laskee, kuinka kauan signaalin kulku satelliitista vastaanottimeen on kestänyt, ja sen perusteella etäisyytensä satelliitista. Kun tiedetään vastaanottimen etäisyys kolmesta satelliitista, voidaan laskea sen sijainti kolmen pallopinnan leikkauspisteessä⁶. Muitakin, parempia tapoja laskemiseen on, mutta niihin ei puututa tässä.

⁶ Leikkauspisteitä on kaksi, mutta yleensä niistä voidaan helposti valita oikea sen perusteella, että tiedetään vastaanottimen olevan maan pinnalla, tai käyttämällä neljännen satelliitin tietoja.

Kolme satelliittia ei kuitenkaan todellisuudessa riitä tarkkaan paikannukseen, sillä GPS-vastaanottimien kellot ovat epätarkkoja. Niin sanotun kellovirheen laskemiseen tarvitaan neljäs satelliitti, jolloin on käytössä neljä yhtälöä neljästä satelliitista ja neljä tuntematonta, kolme paikan koordinaattia ja aika. (Wolf 2004.)



Kuva 1. GPS-järjestelmän osat: 1) kontrolliverkko (eli valvonta-asemat, *monitoring stations*), 2) satelliitit (*space segment*) ja 3) käyttäjät (*user segment*).

Paikan ja ajan laskemiseen tarvitaan siis vähintään neljä satelliittia, ja jos satelliitteja on näkyvissä enemmän kuin neljä, saadaan yleensä parempia tuloksia. Epätarkempia, mutta joissain tapauksissa riittävän tarkkoja, tuloksia voidaan saada myös kolmella satelliitilla, jos oletetaan jokin muuttujista tunnetuiksi. Esimerkiksi voidaan olettaa, että laiva on merenpinnan tasolla ja korkeus siten tunnettu. (Leick 2004, Wolf 2004.)

Satelliiteissa käytetystä tekniikasta aiheutuu useita huomioonotettavia asioita. Satelliittien tilan lähetys signaalissa on olennaista, koska ajoittain satelliitit joutuvat epäkuntoon, jolloin niiden lähettämää tietoa ei pidä käyttää paikannukseen. Satelliitti ei kuitenkaan itse tiedä menneensä epäkuntoon ennen kuin se huomataan ja tieto välitetään satelliitille. Käytetystä tekniikasta on pitkälti kiinni, kuinka nopeasti GPS-vastaanotin kykenee käynnistämisen jälkeen löytämään riittävästi satelliitteja ja keräämään niiltä riittävät tiedot (TTFF, *time to first fix* on aika, joka vastaanottimelta kuluu paikannukseen käynnistyksen jälkeen). Paikan laskemiseen tarvittavat ratatiedot lähetetään satelliitista 30 sekunnin välein (Leick 2004), joten jo tämä tarkoittaa, että ensimmäistä paikannusta käynnistyksen jälkeen ei voi apukeinoja käyttämättä nopeuttaa sitä nopeammaksi. Paikannustarkkuus taas riippuu satelliittijärjestelmää enemmän vastaanottimista ja signaaleja häiritsevistä tekijöistä.

Monet GPS-laitteet määrittävät paikan lisäksi nopeuden. Nopeuden määrittämiseen on useita tapoja. Yksinkertainen tapa, joka toimii kohtuullisen hyvin silloin, kun nopeus pysyy suunnilleen vakiona, on approksimoida käyttäjän sijainnin muutosta ajan kuluessa. Yleisempi tapa on laskea nopeus Dopplerin ilmiön perusteella GPS-satelliitin

lähettämästä signaalista. GPS-laitteiden valmistajat ovat kuitenkin usein vastahakoisia paljastamaan laskentamenetelmiensä yksityiskohtia. (Wolf 2004.)

GPS:n tarkkuus ja sen parantaminen apujärjestelmin

GPS on ollut täysin toiminnassa ja käytettävissä vuodesta 1995, mutta toukokuuhun 2000 asti USA:n hallitus häiritsi tarkoituksella siviilien vastaanottamaa GPS-signaalia niin, että paikannustarkkuus oli vain 30–100 metriä (niin sanottu *selective availability*) (esim. Bonsall ym. 2006). Häirinnän poistamisen jälkeen GPS:n paikannustarkkuus on nykyisin 5-10 metrin luokkaa, kuluttajille myytävien laitteiden valmistajat lupaavat tyyppillisesti esimerkiksi 10 metrin tarkkuuden 90 prosentille havainnoista. Käytännössä tämä saavutetaan hyvissä olosuhteissa, joissa laitteella on yhteys riittävän moneen satelliittiin, joten valmistajien lupauksiin tarkkuudesta ei välttämättä kannata luottaa. Toisaalta ihanneolosuhteissa tarkkuus voi usein olla parempikin, etenkin uusilla herkkillä laitteilla.

Liikkumistutkimuksissa 10 metrin tarkkuus riittänee hyvin ainakin lähtö- ja määrapaikkojen ja summittaisten reittien selvittämiseen. Tarkempaa paikannusta voidaan kaivata, jos esimerkiksi pyritään sovittamaan havaittuja reittejä katuihin tai muihin tunnettuihin elementteihin. Edellä mainitut luvut koskevat lähinnä vaakasuuntaista paikannustarkkuutta, korkeuden tarkkuus on yleensä huonompi (esim. Schuessler ja Axhausen 2008). Liikkumistutkimuksissa korkeuden tarkkuudella ei ole suurta merkitystä, koska korkeustietoa ei yleensä käytetä.

Joissain sovelluksissa, esimerkiksi merenkulussa, ajoneuvojen paikannuksessa ja etenkin kartoituksessa, paikannustarkkuudella on enemmän väliä. Tarkkuutta voidaan tarvittaessa parantaa useilla eri menetelmillä, kuten tunnettuja systemaattisia virheitä korjaavilla apusignaaleilla, nopeuttamalla tiedon saantia tai käyttämällä paikannuksessa apuna jotain muuta tekniikkaa. Seuraavassa esitellään lyhyesti muutama menetelmä. Esimerkiksi maanmittauksen tarpeisiin on kehitetty paljon muitakin menetelmiä.

AGPS (tai A-GPS, englanniksi *assisted GPS*) viittaa joukkoon tapoja, joilla voidaan nopeuttaa ja helpottaa paikannuksen aloittamista. GPS-vastaanottimelle voidaan esimerkiksi toimittaa satelliittien lentoratatiedot etukäteen, tai nopeammalla yhteydellä kuin ne tulisivat satelliitista. Samoin voidaan avustaa vastaanotinta, joka on huonoissa kuuluvuusoloissa. Usein riippuu asiayhteydestä, tarkoittaako AGPS kaikkia näitä avustustekniikoita ylipäätään vai jotain tiettyä tapaa. Myös uusia tapoja kehitetään. (mm. Jarvinen ym. 2002, Särkkä 2007, työssä käytetyn GPS-laitteen käyttöohje.)

Paikannukseen esimerkiksi ilmakehästä aiheutuvaa systemaattista virhettä voidaan korjata maa-asemien avulla, joiden tarkka sijainti tunnetaan. Asemat vastaanottavat satelliittien signaalia, ja laskevat satelliittitiedon perusteella määritetyn sijainnin ja tunnetun sijaintinsa eron. Differentialinen GPS (DGPS) on menetelmä, jossa maa-asemat lähettävät havaitsemansa eron, ja riittävän lähellä olevat vastaanottimet voivat käyttää sitä korjaussignaalina ja paikantaa itsensä tarkemmin. DGPS-signaali on siis alueellista, ja sitä on tarjolla niin ilmaisena kuin maksullisenakin. Differentialista korjausta voidaan joissain tapauksissa soveltaa myös jälkikäteen. Toisessa menetelmässä maa-asemien verkosto lähettää signaalinsa satelliittijärjestelmään, ja satelliitit lähettävät korjaussignaalin käyttäjille. Tällaiset järjestelmät kattavat usein laajemman alueen, esimerkiksi USA:ssa on WAAS, Euroopassa EGNOS ja Japanissa MSAS⁷. Useat kau-

⁷ WAAS = *Wide Area Augmentation System* (USA),
EGNOS = *European Geostationary Navigation Overlay Service* (Eurooppa),
MSAS = *Multi-functional Satellite Augmentation System* (Japani)

palliset vastaanottimet osaavat nykyisin automaattisesti hyödyntää WAAS-, EGNOS- tai MSAS-signaalia tai niitä kaikkia, mutta se ei takaa, että signaalia välttämättä olisi saatavissa. (Särkkä 2007, Maanmittauslaitos 2009, GPS-laitteiden käyttöohjeet.)

GPS:n virhelähteet



Kuva 2. Korkeat rakennukset ja puut häiritsevät usein GPS-signaalin vastaanottamista.

GPS-paikannusta vaikeuttaa kaikki, mikä estää GPS-laitetta saamasta tai säilyttämästä suoraa yhteyttä satelliitteihin, tai heikentää yhteyttä. Paikannus toimii parhaiten, jos satelliitilla on suora ”näköyhteys” satelliitteihin, toisin sanoen pilvetöntä taivasta näkyy mahdollisimman paljon. Tällöinkin ilmakehä ja siinä esiintyvät häiriöt vääristävät signaalia. Yhteyttä heikentävät niin rakennukset (kuva 2), kulkuvälineet kuin puiden lehvästötkin. Etenkin korkeiden rakennusten välissä radiosignaali voi myös heijastua rakennusten seinistä ja muista esteistä ja saapua vasta sitten vastaanottimeen, jolloin signaali kulkee pidemmän matkan kuin satelliitin ja vastaanottimen välillä todellisuudessa on (niin sanottu monitieheijastus, engl. *multi path error*). Lisäksi virhettä aiheutuu vastaanottimista, satelliittien radanmäärittämisestä ja kellovirheistä. (mm. Maanmittauslaitos 2009.)

Usein GPS-paikantimet toimivat huonosti tai eivät saa signaalia ollenkaan sisätiloissa, tunneleissa ja kulkuvälineissä, esimerkiksi junissa

ja busseissa. Tekniikan kehittyminen ja vastaanottimien muuttuminen herkemmiksi pienentää ongelmia jonkin verran etenkin kulkuvälineissä; ainakin busseissa signaalin kuuluvuus on parantunut GPS-avusteisten liikkumistutkimusten alkua ajoista. Henkilöautoissa signaalin vastaanotto toimii jo varsin hyvin, kunhan laite on sijoitettu sopivaan kohtaan. Autonavigaattorien kasvava suosio on tästä hyvä todiste.

Yksi virhelähde on, että käynnistettäessä GPS-vastaanottimelta kuluu jonkin aikaa muodostaa yhteys satelliitteihin, eli löytää taivaalta riittävän monta satelliittia ja vastaanottaa niiden signaalit. Yhteyden muodostamiseen kuluva aika riippuu ainakin käynnistyspaikan olosuhteista, liikkumisnopeudesta yhteydenmuodostuksen aikana (Stopher ym. 2008) ja laitteen ominaisuuksista. Yhteyden muodostamiseen kuluu vähintään useita sekunteja, tyypillisesti puolesta minuutista minuuttiin, ja esimerkiksi liikkuvassa autossa voi kulua useita minuutteja (mm. Stopher ym. 2008 sekä kokeilut kenttätutkimuksen yhteydessä). Edellä mainituilla AGPS-menetelmillä yhteyden muodostamista voidaan kuitenkin nopeuttaa. Yhteyden saaminen on helpointa, jos laite on paikoillaan avoimessa ympäristössä. Liikkuvassa välineessä yhteyttä ei pahimmassa tapauksessa saada ollenkaan.

Yhteyden muodostamiseen kuluva aika riippuu myös siitä, kuinka pitkään vastaanotin on ollut pois päältä, ja siitä, onko sitä siirretty paikasta toiseen sen jälkeen, kun se on viimeksi ollut päällä. Sen mukaan, kuinka paljon käyttökelpoista tietoa laitteella on muistissaan edellisen käyttökerran jäljiltä, puhutaan kylmästä, lämpimästä tai kuumasta käynnistyksestä (suomenkieliset termit eivät liene täysin vakiintuneet, mutta englanniksi käytetään termejä *cold start*, *warm start* ja *hot start*). Eri siru- ja laitevalmistajien

määritelmät sille, mitä mikäkin näistä tarkoittaa, saattavat vaihdella, mutta yleisesti ottaen määritelmät ovat seuraavat:

Taulukko 1. GPS-laitteen käynnistystavat

Käynnistystapa	Muistissa olevat tiedot*	Esimerkki
Kylmäkäynnistys (cold start)	Laitteella ei ole mitään ennakkotietoja.*	Laitetta ei ole käytetty ikinä tai pitkään aikaan, tai sitä on siirretty pitkä matka edellisestä paikasta.
Lämmin käynnistys (warm start)	Laitteen muistissa on tieto edellisestä sijainnista, kellonaika ja almanakkatiedot, mutta satelliittien lentoratatiedot ovat vanhentuneet***.	Laite on ollut pois päältä muutaman tunnin tai yön yli, ja sitten se käynnistetään uudelleen samassa paikassa kuin se sammutettiin.
Kuumakäynnistys (hot start)	Laitteen muistissa on tieto edellisestä sijainnista, kellonaika ja almanakkatiedot, ja niiden lisäksi satelliittien lentoratatiedot ovat voimassa, joten laite tietää mistä etsiä satelliitteja.*	Laite on pois päältä joitakin minutteja, tai ajetaan esimerkiksi tunneliin jossa yhteys katkeaa hetkeksi.

*) Lähde Atmel 2009

**) Riippuu valmistajasta, kuinka monta tuntia vanhat tietojen pitää olla että ne katsotaan vanhentuneiksi.

GPS-signaalia voidaan myös tahallisesti tai tahattomasti häiritä eri tavoin, esimerkiksi muilla signaaleilla (Olsen ja Forssell 2003). Tahaton häirintä on kuitenkin hyvin satunnaista ja tahallinen häirintä toteutetaan usein sotilaallisissa tai vastaavissa yhteyksissä. Myös häirinnän estoa on toteutettu lähinnä sotilaallisiin tarkoituksiin (Laundry ym. 2006). Häirintää ei voitane pitää suurena riskinä liikkumistutkimuksissa edellä mainittuihin virhelähteisiin verrattuna, joten ihmisen aiheuttamaan häirintää tai sen torjumista ei käsitellä tässä sen tarkemmin.

Algoritmien esittelyn yhteydessä (kohta 3.3.1) palataan hieman siihen, miten erilaiset virheet näkyvät liikkumistutkimuksissa kerätyissä aineistoissa, ja miten ongelmia voidaan lieventää tai ratkaista aineistoa käsiteltäessä.

3.1.2 GPS-laitteet liikkumistutkimuksissa

Yleistä

Tässä osiossa kerrotaan, millaisia GPS-laitteita liikkumistutkimuksissa käytetään ja miten laitteet ovat kehittyneet. Lisäksi käsitellään laitteiden hintaa ja tallennusominaisuuksia, sekä käytettävyyttä tutkimusten osallistujien kannalta. Kirjoitettaessa on käytetty hyväksi mainittujen kirjallisuuslähteiden lisäksi myös sopivien GPS-laitteiden etsimistä ja hankkimista varten tehdyssä tutkimustyössä saatuja kokemuksia eri laitteista ja niiden ominaisuuksista.

Laitteen perusominaisuudet

Jotta satelliittien lähettämää GPS-signaalia voidaan hyödyntää, tarvitaan laite, joka osaa vastaanottaa signaalia ja tulkitsemalla sitä paikantaa itsensä. Tällaista laitetta nimitetään GPS-vastaanottimeksi. Liikkumistutkimuksissa tarvitaan GPS-vastaanottimen lisäksi laite, joka tallentaa paikannustiedot riittävän tihein väliajoin. Tällaista laitetta voidaan kutsua vaikkapa GPS-loggeriksi (englanniksi *GPS logger* tai *GPS data logger*). Usein sama laite on sekä GPS-vastaanotin että -loggeri, mutta nimitykset eivät ole täysin vakiintuneita. Siksi tällaisesta liikkumistutkimukseen soveltuvasta laitteesta, joka sekä paikantaa että tallentaa, käytetään tässä työssä yksinkertaisesti nimitystä GPS-laite

(englanniksi usein vain *GPS device*, jolloin asiayhteydestä yleensä selviää, mitä kaikkea laitteella voi tehdä).

Toinen mahdollisuus on siirtää paikannustieto heti tai pienin väliajoin langattomasti GPS-vastaanottimesta tietokantaan, esimerkiksi GPRS-yhteyttä käyttäen (Stopher 2008). Tällöin ei tarvita erillistä tallennuslaitetta, jonka muistiin mahtuisi kaikki kerättävä aineisto, mutta paikannuslaitteen on pystyttävä lähettämään tieto eteenpäin.

GPS-laitteiden historia ja kehitys liikumistutkimuksissa

GPS-laitteiden historia liikumistutkimuksissa alkoi 1990-luvun puolivälin tienoilla USA:ssa. Alussa laitteita käytettiin vain autoissa, koska laitteissa ei ollut sisäistä virtalähdettä, vaan virta otettiin autosta. Sitten on kehitetty akku- ja paristokäyttöisiä, mukana kannettavia GPS-laitteita, joilla voidaan seurata myös joukko- ja kevyen liikenteen käyttäjiä. Akku- ja GPS-tekniikan kehittyessä kannettavat laitteet ovat pienentyneet monen sadan gramman painoisista laatikoista ja erillisistä antenneista taskuun mahtuviksi yksiköiksi, jotka suorittavat useita tehtäviä yhtä aikaa. (Stopher ym. 2008).

Huomattavaa on, että ensimmäiset niin sanotut kannettavat laitteet olivat, lähinnä juuri akkujen takia, niin suuria ja kömpelöitä, että niiden kantamisesta aiheutui käyttäjille huomattavaa vaivaa (kuva 3) (esim. Stopher 2008, Stopher ym. 2008). Vasta kannettavien laitteiden pienentyminen (kuva 4) on mahdollistanut laajemmat GPS:ää käyttävät tutkimukset muillakin kulkutavoilla kuin henkilöautolla kulkevien osalta. Kuvassa 5 on vertailun vuoksi havainnollistettu tässä tutkimuksessa käytetyn laitteen kokoa.

Liikumistutkimuksissa on käytetty sekä varta vasten kehitettyjä että suoraan kaupan hyllyltä saatavia GPS-laitteita (Stopher ym. 2008). Laitekehitys on ollut varsin ripeää GPS:n kaupallisten sovellusten lisääntyessä koko ajan. Laitteiden kehittyminen näkyy paitsi koon pienentymisenä myös virrankulutuksen vähentymisenä, signaalin havaitsemisen parantumisenä ja paikannustarkkuuden parantumisenä. Myös matkapuhelinten kiihaasta kehittämisestä on ollut hyötyä, sillä useat kaupalliset GPS-laitteet käyttävät samanlaisia akkuja kuin matkapuhelimet. Toisaalta matkapuhelinten kehittäjät ovat GPS-sirujen halventuessa integroineet GPS-paikannusta matkapuhelimiin ja tarjoavat sitten matkapuhelimen käyttäjälle paikannukseen perustuvia ohjelmistoja tai palveluja. Lisäksi esimerkiksi Yhdysvalloissa matkapuhelinvalmistajien on huolehdittava siitä, että jos matkapuhelimella soittaa hätänumeroon, soittaja voidaan paikantaa, ja useimmat matkapuhelinvalmistajat hoitavat tämän nimenomaan liittämällä matkapuhelimiin GPS-sirun (Stopher 2008).

Tulevaisuudessa voi olla mahdollista, että esimerkiksi kaikissa matkapuhelimeissa tai kaikissa autoissa on GPS-siru hätätilanteiden, varkauksien ja muiden vastaavien tilanteiden varalta. Tätä automaattista paikannusmahdollisuutta voidaan ehkä hyödyntää myös liikumistutkimuksissa, mutta se ei ole itsestään selvää, kuten esimerkiksi seuraavassa kohdassa Käytettävyys ja myöhemmin kohdassa Yksityisyyden suoja (s. 33) todetaan. Lisäksi kaikilla ihmisillä ei ole matkapuhelinta, ja autossa oleva GPS-laite tallentaa vain automatkat, joten kattavaa otosta ei näin välttämättä saada. GPS-sirujen yleistymistä on luultavasti helpompi hyödyntää hankkeissa, joissa ihmiset voivat vapaaehtoisesti antaa paikantaa itsensä heille sopivana aikana ja siten osallistua sellaisen yleishyödyllisen tiedon tuottamiseen, joka koskee esimerkiksi tiettyä aikaa tai paikkaa⁸.

⁸ Yksi esimerkki tällaisesta hankkeesta on Mobile Millennium, jossa GPS-sirulla varustetun matkapuhelimen omistajat voivat osallistua ajantasaisen liikennetiedon tuottamiseen (<http://traffic.berkeley.edu/theproject.html>).



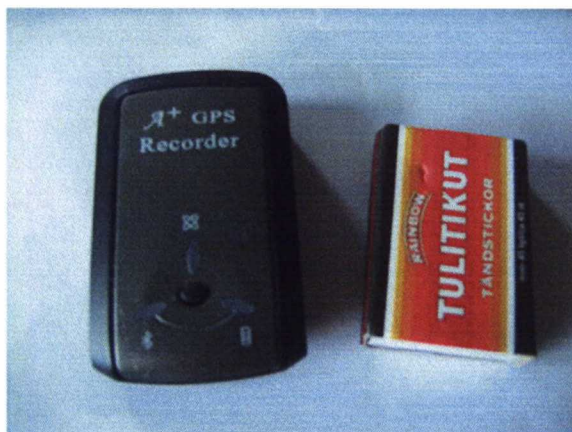
Fig. 1. Early wearable GPS device shown by a mobile telephone.

Kuva 3. Yksi ensimmäisistä liikkumistutkimuksia varten kehitetyistä kannettavista GPS-laitteista, jota käytettiin muun muassa vuonna 2002. Vieressä matkapuhelin koon havainnollistamiseksi. (Lähde: Stopher ym. 2008.)



Fig. 2. New GPS device compared to mobile telephone.

Kuva 4. Seuraavan sukupolven GPS-laite vuodelta 2005. Vieressä matkapuhelin koon havainnollistamiseksi. (Lähde: Stopher ym. 2008.)



Kuva 5. Tutkimuksessa käytetty GPS-laite (iBT747A+) vuodelta 2008. Laitteen ulkomitat ovat 72,2x46,5x20 mm ja paino noin 70 g.

Käytettävyys

Tutkimusten osallistujien kannalta olennaista on, kuinka paljon vaivaa tutkimukseen osallistuminen aiheuttaa. Ei siis ole toivottavaa, että GPS-laitteen kantaminen on hankalaa, tai että sitä joutuu lataamaan useita kertoja päivässä (Stopher ym. 2008). Kun laitteiden herkkyyden lisääntyminen ja tarkkuuden parantuminen ovat tärkeitä tutkittavan tiedon laadun kannalta, on käytettävyys puolestaan vaikuttanut paljon enemmän laitteiden koon pienentyminen ja akkutekniikan kehittyminen.

Stopher ym. (2008) ovat useamman vuoden ajan kehittäneet yhdessä laitevalmistajien kanssa liikkumistutkimuksiin soveltuvaa kannettavaa GPS-laitetta. Myös laitteiden käyttäjien mielipiteitä on kyselty; Stopher ym. (2008) siteeraavat Swannin ja FitzGeraldin tutkimusta, jossa haastateltiin erään tutkimuksen osallistujia, jotka olivat kokeilleet GPS-laitteita. Tutkimuksen mukaan käyttäjän kannalta hyvän laitteen pitäisi mahtua helposti laukkuun tai taskuun, eli olla mahdollisimman pieni ja ohut. Laitteen akun pitäisi kestää mahdollisimman pitkään, mutta vähintään yhden päivän ajan, niin että laitteen lataaminen yöllä riittää. Lisäksi testilaitteessa olleita valoja, jotka ilmoittivat laitteen tilan, ei pidetty intuitiivisina, joten jotkut toivoivat, että laite kertoisi puheäänellä, kun se on laitettu päälle tai saanut yhteyden satelliitteihin. (Stopher ym. 2008.)

Käytettävyys kannalta on myös tärkeää, että laitteen laittaminen päälle ja pois on käyttäjille mahdollisimman vaivatonta. Stopher ym. (2008) raportoivat, että joillakin laitteilla oli taipumus mennä helposti pois päältä, kun ne pudotettiin laukkuun tai taskuun, ja ratkaisuna tähän kehitettiin laite, jonka käyttäjä pystyi laittamaan päälle, mutta ei pois päältä. Uusimpaan, kolmanteen laitesukupolveen saatiin muiden parannusten ohella käyttäjien toivoma laitteen tilan ilmoitus puheäänellä. Lisäksi laitteessa on värinä-tunnistin, jonka avulla laite menee päälle kun se liikkuu ja pois päältä kun se ei liiku, jolloin virtaa säästyy ja turhaa aineistoa kertyy vähemmän. Ratkaisu on toiminut hyvin tehokkaasti. (Stopher ym. 2008.)

Automaattisesti päälle ja pois menevän laitteen haittapuolena on, että jos ihmiset haluavat jättää tallentamatta yksityisiksi katsomiaan matkoja, he eivät voi laittaa laitetta pois päältä kyseisen matkan ajaksi, vaan laite on jätettävä jonnekin. Esimerkiksi Ranskan henkilöliikennetutkimuksessa osallistujille annetaan mahdollisuus laittaa laite pois päältä joidenkin haluamiensa matkojen ajaksi (Marchal ym. 2008).

GPS-laite voi kommunikoida käyttäjälle paitsi valoilla ja äänillä myös monipuolisemmin näytöllä, joka voi näyttää esimerkiksi sijainnin kartalla tai nuolen, joka osoittaa annettua määränpäättä kohti. Usein GPS-tutkimuksissa on kuitenkin käytetty mahdollisimman yksinkertaisia laitteita, joissa ei ole näyttöä. Marchal ym. (2008) perustelevat tätä liikenneturvallisuudella ja sillä, että näyttö saattaisi vaikuttaa matkustuskäyttäytymiseen. Liikenneturvallisuuden takaamiseksi on varmistettava paitsi se, että osallistujilla ei ole intressiä kiinnittää ylimääräistä huomiota GPS-laitteeseen ajaessaan, myös se, että laitteen voi kiinnittää tai asettaa ajoneuvossa sellaiseen paikkaan, ettei se lähde liikkeelle ajon aikana.

Useimmat matkapuhelimen käyttäjät ovat tottuneet huolehtimaan siitä, että matkapuhelin on koko ajan käyttövalmis. Matkapuhelimeen integroitu GPS-siru vaikuttaa siis ensi kuulemalta ideaaliselta tavalta saada tutkimukseen osallistuva käyttäjä pitämään paikannuslaite koko ajan vaivattomasti mukana ja päällä ja muistamaan laitteen lataaminen. Huono puoli on, niin matkapuhelimiin kuin kämmentietokoneisiinkin yhdistetyn GPS:n kohdalla, että yhdistelmälaitteiden virrankulutus on usein suhteessa suurempaa kuin yksittäisten laitteiden, ja laitetta voi joutua lataamaan aina kun vain on mahdollista, mikä aiheuttaa suurta vaivaa osallistujille (Bonsall ym. 2006). Akkutekniikan kehittyminen saattaa tulevaisuudessa

ratkaista latausongelman. Mutta jos oletetaan, että tutkimuksissa voidaan käyttää ihmisten omia matkapuhelimia, voi kestää kauankin, ennen kuin riittävän kattavalla osalla väestöstä on tarkoituksen sopiva matkapuhelin ja siinä riittävän pitkään kestävä akku.

Hinta

GPS-laitteiden hinnat ovat viime vuosina halventuneet, mutta jos tarkoituksena on toteuttaa laajoja tutkimuksia, myös hinnalla on merkitystä. Käytännössä on optimoitava laitteiden määrän, otoksen koon ja tutkimuksen keston suhteita. Laajoja tutkimuksia voidaan toteuttaa pienemmällä määrällä laitteita, jos aikaa on paljon. On kuitenkin huomioitava, että varsinaisen tutkimusajan lisäksi aikaa kuluu laitteiden toimittamiseen osallistujille ja keräämiseen takaisin sekä tiedon purkamiseen laitteilta. Hintaan vaikuttavat myös ominaisuudet, joita laitteilta vaaditaan. Monipuolisemmilla laitteilla voidaan saada parempaa aineistoa, jonka jatkokäsittely on nopeampaa ja helpompaa. Tarvittavaan aineiston laatuun vaikuttaa myös se, miten aineistoa on tarkoitus jatkokäsitellä ja käyttää. Käyttötarkoitus vaikuttaa myös GPS-osuudesta aiheutuviin kokonaiskustannuksiin. (Bonsall ym. 2006.)

Tallennettava aineisto

Kuluttajille myytävät GPS-laitteet tallentavat usein päivämäärän ja kellonajan, pituus- ja leveyskoordinaatit, korkeuden ja nopeuden (kuva 6). Jotkin laitteet tallentavat näiden lisäksi paikannustarkkuudesta kertovia tietoja (kuten näkyvissä ollut satelliittien määrä tai HDOP⁹) sekä suunnan (engl. *heading*) (mm. Stopher ym. 2008). Usein on mahdollista valita, mitä tietoja tallennetaan. Mitä enemmän tietoja tallennetaan, sitä vähemmän pisteitä muistiin luonnollisesti mahtuu. Jos tietoa on enemmän, jatkokäsittely voi kuitenkin olla helpompaa, joten tässäkin on tehtävä valintoja sen mukaan, miten ja mihin aineistoa lopulta käytetään.

INDEX	DATE	TIME	LATITUDE	N/S	LONGITUDE	E/W	ALTITUDE	SPEED
1	18.2.2009	6:25:04	60.218957	N	24.811325	E	33.909477	18.879990
2	18.2.2009	6:25:05	60.219136	N	24.811498	E	49.289238	19.360521
3	18.2.2009	6:25:06	60.219243	N	24.811624	E	60.089798	19.364946
4	18.2.2009	6:25:07	60.219343	N	24.811761	E	67.597839	19.168406
5	18.2.2009	6:25:08	60.219414	N	24.811885	E	73.098648	19.236557
6	18.2.2009	6:25:09	60.219487	N	24.812009	E	78.653877	19.676012
7	18.2.2009	6:25:10	60.219523	N	24.812126	E	80.929131	20.497009
8	18.2.2009	6:25:11	60.219572	N	24.812253	E	84.887024	21.111141
9	18.2.2009	6:25:12	60.219604	N	24.812381	E	86.809120	22.159973
10	18.2.2009	6:25:13	60.219623	N	24.812490	E	87.170433	20.926601
11	18.2.2009	6:25:14	60.219659	N	24.812624	E	89.315422	22.467491
12	18.2.2009	6:25:15	60.219674	N	24.812753	E	90.571518	24.393473
13	18.2.2009	6:25:16	60.219692	N	24.812888	E	91.778130	25.017380
14	18.2.2009	6:25:17	60.219710	N	24.813032	E	92.340546	28.468441
15	18.2.2009	6:25:18	60.219721	N	24.813183	E	92.669434	27.649670

Kuva 6. Esimerkki tutkimuksessa käytetyn GPS-laitteen (iBT747A+) tallentamista tiedoista. (Vasemmalta oikealle: tietueen järjestysnumero, päivämäärä (UTC), UTC-kellonaika, leveysaste, pallonpuolisko (pohjoinen N / eteläinen S), pituusaste, pallonpuolisko (itäinen E / läntinen W), korkeusasema ja nopeus.)

Nykyisillä suurilla laskentakapasiteeteilla suurikin tietoaaineisto voidaan automaattisesti käsitellä ja muuntaa sopivaan muotoon valmiita jatkokäsittelyohjelmistoja, esimerkiksi

⁹ HDOP = horizontal dilution of precision. Luku kertoo, millä tavoin satelliitit ovat levittäytyneet taivaalle GPS-laitteesta katsottuna, eli kuinka luotettava mittaus on. Jos HDOP-luku on suuri, tulos on epäluotettava. Jos luku on esimerkiksi alle 3, niin mittaustulosta voidaan pitää luotettavana. (Stopher ym. 2008.)

paikkatieto-ohjelmia ja tulkinta-algoritmeja, varten. Jotkin laitteet tuottavat laitetoimittajan oman ohjelmiston mukaisia tiedostoja ja lisäksi standardinmukaisia¹⁰ tiedostoja, mutta standardinmukaisessa muodossa ei välttämättä ole aivan kaikkia kerättyjä tietoja mukana. Tiedonkäsittelyä helpottaa, jos GPS-laite antaa ulos tietoa standardinmukaisessa tai muussa helposti käsiteltävässä muodossa (esimerkiksi tekstimuotoisena pilkuilla erotettuna csv- eli *comma separated values* -muodossa). Tällöin voidaan käyttää useita eri tallennuslaitteita samassa tutkimuksessa, jos käytetään vain niitä tietoja, jotka kaikki laitteet ovat tallentaneet.

Tallennuskapasiteetti

GPS-laitteiden tallennuskapasiteetti, eli se kuinka monta reittipistettä laitteen muistiin voidaan tallentaa, vaihtelee huomattavasti. Esimerkiksi kuluttajille retkeilykäyttöön tarkoitettuihin laitteisiin voi mahtua 1 000 reittipistettä, mikä ei riitä alkuunkaan liikkumistutkimuksissa. Sen sijaan ihmisten tai ajoneuvojen seuraamiseen tarkoitettut laitteet soveltuvat tarkoitukseen paremmin, sillä niihin voi mahtua kymmeniä tai satoja tuhansia reittipisteitä. Tallennettavien pisteiden määrä ei välttämättä näy laitteen hinnassa, vaan enemmänkin on kyse siitä, mihin käyttöön laite on tarkoitettu ja miten paljon muistia siinä on.

Helpoin tapa saada sama tallennuskapasiteetti riittämään pidempään on tietenkin harventaa tallennusväliä. Esimerkiksi viiden tai kymmenen sekunnin välein, tai vieläkin harvemmin, tallentaminen riittää usein aivan hyvin. Joissakin laitteissa taas on mahdollisuus tallentaa vain silloin, kun nopeus ylittää annetun rajan. Esimerkiksi Stopher ym. (2008) ovat todenneet, että jos tallennetaan vain silloin, kun nopeus on yli 3 km/h, turhaa aineistoa tallentuu huomattavasti vähemmän. Tällöin voidaan tosin menettää osa informaatiosta. Kolmas mahdollisuus on käyttää älykästä, liikkeen tunnistavaa (esimerkiksi aiemmin mainitun kaltaista tärinätunnistimella varustettua) GPS-laitetta, joka tallentaa vain silloin kun se todella on liikkeessä. Näin tallennuskapasiteetin tarvetta voidaan vähentää menettämättä informaatiota hitaasta liikkeestä, esimerkiksi liikkeellelähdoistä ja jarrutuksista.

Tallennuskapasiteetin tarve riippuu laitteen käyttötarkoituksesta eli tutkimuksen tavoitteista ja metodeista. Jos esimerkiksi halutaan tallentaa henkilön liikkumista yhden sekunnin välein, kertyy pisteitä 3 600 tunnissa. Jos oletetaan, että ihminen on poissa kotoaan 16 tuntia päivässä ja häntä paikannetaan koko tämä aika, kertyy pisteitä jo yhden päivän aikana 57 600. GPS-paikannuksen epätarkkuudesta johtuen paikoillaan oleva GPS-laite ei yleensä ymmärrä olevansa paikoillaan, vaan laskettu sijainti vaihtelee muutaman kymmenen metrin säteellä (Stopher 2008). Pisteitä siis tallentuu myös laitteen ollessa paikoillaan. Kuitenkin Stopherin (2008) mukaan ihminen yleensä käyttää liikkumiseen noin 70 minuuttia päivässä. Suomessakin tilanne näyttää olevan suunnilleen sama, sillä vuosien 2004–2005 valtakunnallisessa henkilöliikenne-tutkimuksessa saatu keskimääräinen kokonaismatka-aika on 70,8 minuuttia henkilöä kohti vuorokaudessa¹¹. Toki on otettava huomioon, että 70 minuuttia on vain keskiarvo, mutta joka tapauksessa keskimääräinen tallennustarve vähenee huomattavasti edellä oletetusta 16 tunnista vuorokautta kohti, jos pystytään käyttämään jotain edellä kuvatuista tekniikoista.

¹⁰ NMEA-standardi säätelee, missä muodossa GPS-laitteet välittävät tietoa.

¹¹ Tarkka lukuarvo ei ole tässä yhteydessä oleellinen, mutta tarkemmin sanottuna kokonaismatka-aika (min/hlö/vrk) on 70,8 minuuttia kotimaanmatkoja, 4,7 minuuttia ulkomaanmatkoja, tai 75,7 minuuttia kaikkia matkoja (Henkilöliikennetutkimus 2004–2005; <http://www.hlt.fi/index.htm> -> Tulokset -> Yleiskuva liikkumisesta -> <http://www.hlt.fi/tulokset/4.xls>).

3.2 GPS:n käyttö liikkumistutkimuksissa

3.2.1 Paikannus ja muut uudet tekniikat liikkumisen tutkimisessa

Yleistä

Uusien tekniikoiden käyttämistä muun muassa liikkumistutkimuksissa on tutkittu jo vuosia. Vuonna 2006 valmistui tutkimus (Bonsall ym. 2006), jossa on kattavasti vertailtu erilaisia uusia tekniikoita ja niiden soveltuvuutta (Ison-Britannian) henkilöliikennetutkimukseen. Parissa vuodessa tieto on ehtinyt osittain vanhentua, mutta tiedot muun muassa eri tekniikoiden kustannustekijöistä ja soveltuvuudesta henkilöliikennetutkimuksiin sekä esitetyt tulevaisuudennäkymät ovat edelleen ajankohtaisia.

Seuraavaksi käsitellään tarkemmin muun muassa syitä elektronisten apuvälineiden käyttöönottoon liikkumistutkimuksissa, sitä mitä tietoja GPS:llä voi ja ei voi kerätä, muita paikannusmenetelmiä, muiden alojen sovellusten hyödyntämistä liikkumistutkimuksissa, GPS:n kustannuksia ja yksityisyyden suojaa. Kysymyksiä on tässä lähestytty pääosin henkilöliikennetutkimusten näkökulmasta, mutta useat huomiot pätevät yhtä hyvin mihin tahansa liikkumistutkimuksiin. Koko ajan on syytä muistaa, että paikannuksen ja muiden uusien tekniikoiden ei tarvitse täysin korvata perinteisiä tutkimusmenetelmiä, vaan ne voivat etenkin aluksi olla täydentäviä menetelmiä. Tällöin on kuitenkin todennäköistä, että vastaajiin kohdistuva kuormitus ja tutkimusten kustannukset ovat korkeammat kuin silloin, jos käytettäisiin vain yhtä menetelmää. Uudet tekniikat tuovat mukanaan myös täysin uutta tutkimusta, kun päästään tutkimaan asioita, joiden tutkiminen vanhoin menetelmin on ollut liian työlästä tai jopa mahdotonta.

Henkilöliikennetutkimusten nykytila

Henkilöliikennetutkimuksissa on perinteisesti käytetty menetelmänä paperisen matkapäiväkirjan täyttämistä, ja matkapäiväkirja on joko postitettu tutkijoille, tai sen tiedot on kysytty puhelin- tai henkilökohtaisella haastattelulla (Murakami ja Wagner 1999). Perinteisiin matkapäiväkirjatutkimuksiin liittyy kuitenkin useita ongelmia, joita käydään seuraavassa lyhyesti läpi siltä osin kuin GPS:n toivotaan niitä ratkaisevan. Perinpohjaisemmin ongelmia ovat käsitelleet esimerkiksi Stopher ja Greaves (2007). Ongelmien ratkaisemiseksi on kehitetty monia erilaisia uusia keinoja ja apuvälineitä, paikannuksen lisäksi esimerkkejä ovat internet-kyselyt ja mukana kannettavat elektroniset matkapäiväkirjat sekä näiden yhdistelmät (esim. Bonsall ym. 2006).

Perinteisten menetelmien ongelmia on käsitelty useissa lähteissä; ongelmat ovat useimmiten samoja eri puolilla maailmaa. Esimerkiksi Guensler ja Wolf (1999) listaavat joitakin perinteisen matkapäiväkirjan ongelmia, joihin toivotaan saatavan helpotusta GPS-laitteen ja elektronisen matkapäiväkirjan yhdistelmästä: 1) Ihmisten muistin epätarkkuudesta aiheutuu väärää ja puutteellista matkojen raportointia, usein jopa kokonaisia matkoja unohtuu. 2) Paperisen matkapäiväkirjan täyttäminen on aikaa vievää, joten osallistujat väsyvät usein täyttämään sitä, minkä seurauksena useamman päivän matkojen kerääminen on haastavaa, eikä matkoja yleensä kerätä kuin yhdeltä tai kolmelta päivältä. 3) Reittiä ei yleensä saada selville, koska sen raportointi olisi niin työlästä. Lisäksi on huomattu, että ne matkat, jotka unohdetaan kokonaan ilmoittaa, ovat usein lyhyitä (muun muassa Marchal ym. 2008).

Suurimpia ongelmia ovat siis kerätyn tiedon epätarkkuus ja vastaamisesta osallistujille aiheutuva kuormitus. Tiedon epätarkkuus ja matkojen puuttuminen on ongelmallista, koska liikkumistutkimusten tuloksia käytetään liikennemallien ja -ennusteiden

tekemiseen ja niitä puolestaan liikennesuunnitteluun. Ennustettavia asioita ovat esimerkiksi matkojen määrä päivässä asukasta kohti matkaryhmittäin, matkojen lähtö- ja määräpaikat sekä matkojen keston jakaumat ja matkojen ajallinen jakautuminen (Guensler ja Wolf 1999), joten olisi toivottavaa, että niistä saadaan riittävän tarkkaa tietoa tutkimuksissa. Vastaamisen helpottumisen toivotaan sekä mahdollistavan pidemmän ajanjakson tietojen keräämisen että kasvattavan vastausprosentteja, jotka ovat pienentyneet aikojen saatossa; samalla myös vastausprosenttien pienenemisestä johtuva kustannusten nousu halutaan kuriin (Bricka 2008).

Eräs erityinen tiedon epätarkkuuden osa-alue on haastatteluissa saatujen osoitteiden epätarkkuus. Lähtö- ja määräpaikkojen sijainnit jollain tarkkuudella tarvitaan liikenteen mallinnukseen. Ongelma on merkittävä, koska Stopherin (2004) arvion mukaan ihmiset tietävät kotiosoitteensa, ja jotkut muistavat työpaikkansa osoitteen, mutta useimpia muita osoitteita ei muisteta tai tiedetä. Ihmiset siis tietävät kyllä missä he käyvät, mutta usein he eivät tiedä paikkojen tarkkoja osoitteita. Perinteisesti sijaintien välittäminen haastatteluissa muulla tavoin kuin osoitteina on vaikeaa, koska haastattelijat voivat olla toisella puolella maata eikä heillä voi olla paikallistuntemusta koko maan alueelta. Osoitteiden geokoodaamisessa maantieteelliseksi koordinaateiksi on myös omat haasteensa, jotka johtuvat paitsi vastaajien antamien osoitteiden epätarkkuuksista ja virheistä myös geokoodausohjelmistojen ja -tietokantojen rajoituksista (Greaves 2004).

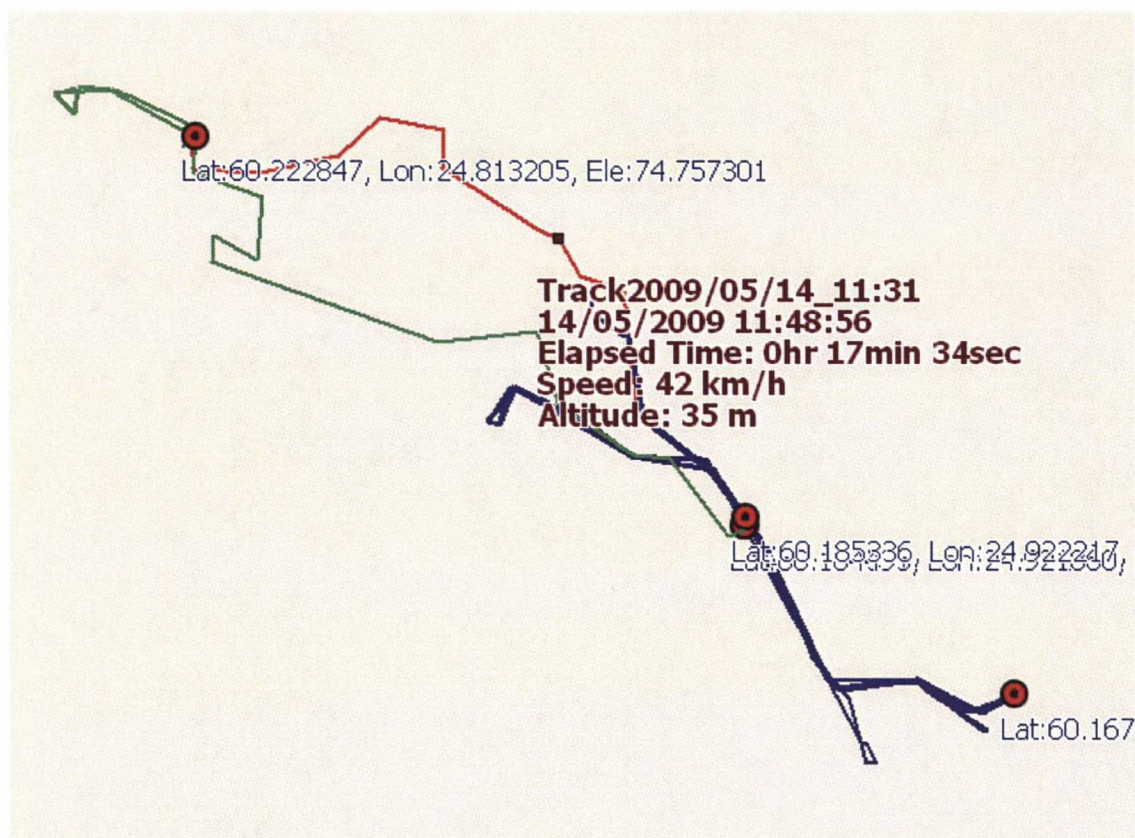
Helpoin ratkaisu sijainninmääritysongelmaan on hankkia vastaajilta osoitteen sijaan suoraan koordinaatit. Yksi ratkaisu on, että vastaajat voivat osoittaa sijainnin suoraan kartalla, jos osaavat. Näin voidaan tehdä henkilökohtaisissa haastatteluissa ja etenkin internet-haastatteluissa. Puhelinhaastatteluissa taas haastattelijoiden avuksi on kehitetty erilaisia tietokantoja, joista he voivat etsiä esimerkiksi paikkoja tai liikkeitä, joissa vastaaja on käynyt (Greaves 2004). Kaikkia paikkoja ei kuitenkaan välttämättä osata näyttää edes kartalta, ja toisaalta internet-kyselyihin saadaan vastaus vain osalta väestöstä.

Henkilöliikennetutkimuksissa GPS-laitteita on tähän mennessä käytetty lähinnä perinteisillä tavoilla kerättyjen tietojen validointiin eikä vielä suuressa määrin varsinaiseen tietojen keräämiseen. Validointi tehdään yleensä toteuttamalla varsinaisen henkilöliikennetutkimuksen yhteydessä GPS-osio (engl. usein *GPS component*, tai GPS-osioon valittuun otokseen viittaava *GPS subsample* tai *GPS sample*). Tällöin osalle osallistujista annetaan GPS-laite kannettavaksi sen lisäksi, että he osallistuvat perinteiseen tutkimukseen. Vertaamalla GPS-osion osallistujien matkapäiväkirjoja GPS-aineistoon voidaan yrittää selvittää, kuinka suuri osa matkoista jää raportoimatta perinteisessä tutkimuksessa, millaiset matkat jäävät raportoimatta ja millaiset ihmiset matkoja unohtavat (Wolf ja Lee 2008). Matkojen määrien lisäksi voidaan verrata esimerkiksi päiväkirjassa ilmoitettuja ja GPS:n tallentamia matkojen pituuksia ja kestoja, ja saaduilla tiedoilla voidaan haluttaessa korjata varsinaisen otoksen tuloksia.

Mitä tietoja GPS:llä voi ja ei voi kerätä

Usein lähdetään liikkeelle nykyisin kerättävästä tiedosta ja mietitään, miten se saataisiin kerättyä paremmin tai helpommin. Perusteena tälle on muun muassa vertailtavuuden säilyttäminen edellisiin tutkimuksiin nähden. On siis paikallaan tarkastella, mitä nykyisin kysyttävistä tiedoista GPS:llä voidaan kerätä ja mitä ei, ja millaisilla täydentävillä menetelmillä loput tiedot voitaisiin kerätä.

Kuten edellä todettiin, GPS-laite tallentaa oleellisesti reittipisteitä ja niihin liittyviä tietoja (kuva 7). GPS:llä on siis mahdollista kerätä tiedot kuljetusta reitistä (reittipisteiden pituus- ja leveyskoordinaatit) ja reittipisteisiin liittyvistä päivämääristä, kellonajoista, nopeuksista ja korkeusasemista (Stopher 2004).



Kuva 7. GPS:llä kerättävät tiedot. Kuvassa näkyy kuljettu reitti (ja joidenkin pisteiden koordinaatit), ja tietyssä pisteessä tiedoston nimi, päivämäärä, kellonaika, matkan alusta kulunut aika, nopeus ja korkeusasema.

GPS:llä ei kuitenkaan pystytä, ainakaan suoraan, keräämään muita henkilöliikennetutkimuksissa usein kysyttäviä tietoja, kuten kulkutapaa, matkan tarkoitusta, mukana olevien henkilöiden määrää ja matkakohdetta tai matkan kustannuksia (esimerkiksi joukkoliikennelipun hinta tai pysäköintikustannukset) (Stopher 2004). GPS:llä ei myöskään saada näkyviin esimerkiksi terveyteen liittyvissä tutkimuksissa tarvittavia aktiivisuus- tai energiankulutuslukemia.

Henkilöiden määrän, tarkan matkakohde ja kustannusten selvittäminen vaativat edelleen, että matkan tekijä kirjaa tiedon ylös. Matkan kulkutapa ja tarkoitus on mahdollista päätellä jollain tarkkuudella, etenkin jos otetaan avuksi paikkatietoaineistot; aktiivisuustietoja taas voidaan kerätä siihen tarkoitukseen tehdyillä laitteilla, jos tutkittavat kantavat niitä mukanaan GPS-laitteen lisäksi (Bonsall ym. 2006).

Matkojen lähtö- ja saapumisajat ja -paikat ovat hieman harmaata aluetta. GPS-laite ei sinänsä osaa erottaa, missä yksi matka päättyy ja toinen alkaa. Tallennettu aineisto on kuitenkin mahdollista jakaa matkoiksi joko käsin tai automaattisesti algoritmeja ja päättelysääntöjä hyödyntäen. Kun yhtenäinen aineisto on saatu muutettua matkoiksi, saadaan helposti matkojen lähtö- ja saapumisajat ja -paikat sekä matkojen kestot, ja matkojen pituudet voidaan laskea reittitiedon perusteella. (Bonsall ym. 2006.)

Koska yksi tavoite tiedon keräämisessä GPS:llä on saada tarkempaa tietoa kuin aiemmin, kannattaa kiinnittää huomiota myös siihen, mitä tietoja itse asiassa pystytään keräämään nykyistä tarkemmin. Stopherin (2008) mukaan GPS:llä matkojen lähtö- ja määräpaikoista voidaan saada nykyistä tarkempaa tietoa (etenkin koska ihmiset usein eivät ole selvillä matkakohteidensa osoitteista), samoin lähtö- ja saapumisajoista ja matkojen kestoista ja pituuksista.

Paikkatietoaineistoja hyödyntämällä voidaan myös päätellä GPS-aineistosta tietoja, joita tutkimuksissa ei aikaisemmin ole kysytty, koska se on koettu vastaajille liian työlääksi tai vaikeaksi, hyvänä esimerkkinä kuljettu reitti. Samoin esimerkiksi kuljetun matkan jakaminen erityyppisille teille ja kaduille on helppoa, jos reitti sovitetaan paikkatietoaineistoon, jossa nämä tiedot ovat valmiina; myös matkanopeudesta ja sen vaihtelusta saadaan tarkkaa tietoa (Murakami ja Wagner 1999). Paikkatietoaineistojen kehittyessä voitaisiin tarkastella matkan sujuvuutta tutkimalla matkanopeutta ja sen vaihteluita ja vertaamalla niitä nopeusrajoitukseen tai tavoitenopeuteen. Useiden tutkittavien nopeustietoja yhdistämällä voitaisiin saada kuva liikenneverkon palvelutasosta eli esimerkiksi ruuhkaisimmat kohdat näkyviin, ja miksei samankaltaista analyysiä myös joukkoliikenteen osalta, jos otoskoko on riittävä (Stopher 2008).

Osoitteiden geokoodaamiseen liittyvän ongelman GPS-laitteiden käyttäminen ratkaisee sikäli, että GPS-aineistosta saadaan suoraan koordinaatit (ongelmaksi jää enää päätellä, mitkä ovat matkan lähtö- ja määräpaikat; tähän palataan kappaleessa 3.3.1) Koordinaatit saadaan kuitenkin vain niin tarkasti kuin GPS-signaali missäkin paikassa kuuluu, joten täysin ongelmaton ei GPS-tutkimuskaan nykYTEKNIKALLA tässä suhteessa ole.

Taulukkoon 2 on vielä koottu yhteenveto siitä, mitä tietoa GPS:llä pystytään ja ei pystytä keräämään. Taulukko on hieman mukaillen käännetty Bonsallin ym. (2006) taulukosta 1.

Paikannustiedon yhdistäminen muihin tiedonkeruumenetelmiin

GPS:llä voidaan kerätä vain osa halutusta tiedosta, ja liikkumistutkimukset koostuvat useista osista, joista GPS-osuus on vain yksi. GPS-aineistoa joudutaan täydentämään vähintäänkin sosiodemograafisilla ja muilla taustatiedoilla. Mahdollisesti joudutaan keräämään normaalia enemmän taustatietoja, joita käytetään apuna puuttuvien matkaan liittyvien tietojen kuten kulkutapojen ja matkan tarkoitusten päättelyssä. Näiden tietojen tarkkuuteen GPS ei luonnollisesti vaikuta, eli taustatiedoissa tulee olemaan samat puutteet kuin tähänkin asti.

Kerättyä GPS-aineistoa voidaan käyttää sellaisenaan, mutta on myös mahdollista käyttää sitä esimerkiksi puhelimitse tai internetissä tehtävän tutkimuksen apuna niin, että GPS-aineisto näytetään käyttäjälle jälkikäteen (engl. *prompted recall*) ja kysytään samalla tarkempia tietoja matkasta tai liikkumisesta (esim. Marchal ym. 2008 ja Bricka 2008). Tämä on vastaajalle ehkä vähemmän kuormittavaa kuin matkapäiväkirjan täyttäminen, mutta kuormittavampaa kuin se, että hänen ei laitteen kantamisen lisäksi tarvitse mitenkään muuten antaa tietoja yksittäisistä matkoista. Jos osallistuja on pitänyt GPS-laitetta mukana useita päiviä, tarkemmat tiedot voidaan myös kysyä esimerkiksi yhden päivän matkoista, ja käyttää näitä tietoja hyväksi kun päätellään lisätietoja muiden päivien matkoista.

Yksi mahdollisuus saada kuhunkin matkaan liittyvät tiedot, joita GPS:llä ei voida kerätä, on yhdistää GPS-paikannusominaisuus kämmentietokoneeseen tai muuhun laitteeseen, siis elektroniseen matkapäiväkirjaan, johon käyttäjä syöttää tietoja matkasta matkan alkaessa tai matkan aikana. Esimerkiksi ajoneuvojen päästöjen tutkimiseen on käytetty myös menetelmiä, joissa GPS:n ja elektronisen matkapäiväkirjan lisäksi ajoneuvoon on asennettu mittalaitteisto, jolla saadaan automaattisesti selville muun muassa ajoneuvon paikka, nopeus, polttoaineenkulutus ja siten päästöt joiltain osin (Guensler ja Wolf 1999).

Taulukko 2. Yhteenveto tiedoista, joita GPS-laitteen avulla voidaan ja ei voida kerätä. (Lähde: Bonsall ym. (2006), Table 1.) Taulukon loppuun on tehty joitain huomautuksia (*-*), joita ei ole alkuperäisessä lähteessä.**

Tietolaji	Voidaanko tieto kerätä GPS:llä
Matkan tarkoitus	Kyllä, epäsuorasti päättelemällä mm. taustatietojen ja paikkatietoaineistojen avulla.
Lähtöaika	Kyllä, päätellään sijainnin muutoksista ja liikkeestä.
Saapumisaika	Kyllä, päätellään mm. pysähtymisestä.
Lähtöpaikka (kunta/kaupunginosa)	Kyllä, koordinaateista joissa ollaan lähtöajankohtana.
Saapumispaikka (kunta/kaupunginosa)	Kyllä, koordinaateista joissa ollaan saapumisajankohtana.
Kulkevat (kävelymatkoista vain 1 mailia pidemmät tai yli 20 minuuttia kestäneet)	Kyllä, epäsuorasti päättelemällä nopeusprofiilien ja reitti- ja pysäkkietojen avulla.
Kuljettu matka	Kyllä, voidaan laskea koordinaateista tai sovittamalla reitti (tie)verkolle.
Matkustajien määrä (jaoteltuna lapsiin ja aikuisiin)	Ei, jos käytetään ajoneuvoihin asennettavia laitteita. Kyllä perheenjäsenten osalta, jos kaikki perheenjäsenet käyttävät omaa kannettavaa laitetta.*
Matka-aika	Kyllä, automaattisesti ja tarkasti.
Lipputyypit joukkoliikenteessä	Ei.
Matkan kustannus	Ei.
Nousujen määrä (montaako liikennevälinettä matkalla käytettiin)	Kyllä, epäsuorasti päättelemällä tunnistetuista kulkutavoista ja kulkutavanvaihtoista.
Mitä autoa, moottoripyörää jne. käytettiin (jos taloudessa on useita)	Kyllä, jos käytetään ajoneuvoon asennettavia laitteita Ei, jos käytetään kannettavia laitteita.
Matkustaja vai kuljettaja (lähinnä jos käytettiin henkilöautoa)	Ei.
Kuljettajat: minne pysäköitiin ja paljonko pysäköinti maksoi	Pysäköintipaikka kyllä (jos käytettävissä on riittävän tarkkaa paikkatietoa pysäköintipaikoista), kustannusta ei.
Kuljettajat: maksetut tietullit ja ruuhkamaksut yms.	Todennäköisesti ei. Jos käytettävissä on paikkatietokanta maksuperusteista, päättely saattaa olla mahdollista.
Matkan lähtö- ja määräpaikka postinumeron tarkkuudella	Kyllä, postinumeron tarkkuudella ja vieläkin tarkemmin, sillä tarkkuudella kuin GPS toimii.**
Kaikki noin 50 metriä pidemmät kävelymatkat	Kyllä, mutta kylmäkäynnistyksen hitauden vuoksi lyhyitä matkoja saatetaan menettää kokonaan.
Lapset ja nuoret: ulkona vietetty aika, jota ei lasketa matkaksi (esimerkiksi leikkiminen, rullalautailu, pyöräily)	Kyllä, jos käytetään kannettavaa laitetta.***

*) Muiden kuin perheenjäsenten osalta tietoa ei tietenkään saada.

**) Jopa katuosoite voi hyvissä oloissa olla mahdollista päätellä.

***) Voi olla vaikeaa erottaa oleskelua varsinaisista matkoista.

Muut paikannustekniikat

Satelliittipaikannuksen lisäksi liikkumistutkimuksissa on mahdollista soveltaa myös muita paikannustekniikoita. Esimerkiksi matkapuhelinten paikantamista tukiasemien avulla on käytetty (esimerkiksi Krygsman ym. 2008). Stopher (2008) mainitsee myös Wi-Fi-paikannuksen ja aktiivisiin Radio Frequency Identification (RFID) -tunnisteisiin perustuvan paikannuksen mahdollisina tulevaisuuden sovelluksina. Mahdollisia ovat

myös sovellukset, joissa käytetään useampia paikannustekniikoita toistensa täydentämiseen. Tällöin voidaan käyttää toista menetelmää apuna niillä alueilla, jotka varsinaisella tekniikalla jäävät katveeseen.

Matkapuhelinten tukiasemia voidaan käyttää yhdessä GPS:n kanssa paikannukseen esimerkiksi sisätiloissa ja tunneleissa. Toinen keino jatkaa paikannusta, kun GPS-signaali häviää, ovat GPS/INS-tekniikat, joissa INS tulee sanoista *Inertial Navigation System*, suomeksi voidaan puhua inertiapaikannuksesta. INS voi esimerkiksi kiihtyvyysanturien avulla jatkaa paikannusta lyhyen aikaa. (mm. Särkkä 2007.)

Matkapuhelinverkkojen tukiasemiin perustuvan paikannuksen tarkkuus on lähtökohtaisesti paljon huonompi kuin GPS:n ja vaihtelee sen mukaan, miten tiheässä tukiasemia on. Kaupungeissa tarkkuus on satojen metrien, keskusta-alueilla jopa kymmenien metrien luokkaa, kun taas harvaanasutuilla alueilla tukiasemia on harvassa ja tarkkuus voi olla kilometrien luokkaa. Huonomman tarkkuuden ei välttämättä tarvitsisi olla ongelma, jos liikenne-ennusteita halutaan tehdä melko suurten osaluueiden välille, mutta tukiasemien kattamien alueiden rajat eivät useinkaan ole yhteensopivia liikenne-ennustealueiden rajojen kanssa. (Krygsman ym. 2008.)

Matkapuhelinpaikannus voi kuitenkin olla kannattavaa alueilla, joissa tietoa ennestään on vähän, eikä tarvita kovin tarkkaa tietoa. Koska matkapuhelimet usein ovat tutkittavien omia, niiden hankkimisesta ei aiheudu kustannuksia, mutta toisaalta silloin voidaan tutkia vain matkapuhelinten omistajia, mikä vääristää otosta. (Krygsman ym. 2008.)

Kustannukset

Harkittaessa uusien tekniikoiden käyttöönottoa yksi tärkeä tekijä on hinta. Ei riitä, että uusi tekniikka toimii, vaan ollakseen kannattavaa sen pitäisi joko tuoda kustannussäästöjä tai tuottaa entistä parempaa (tarkempaa, tarkoituksenmukaisempaa tai kokonaan uutta) tietoa. Uusien tekniikoiden, myös GPS:n, osalta hinta muodostuu useista eri tekijöistä, joista osa on samoja kuin perinteisissä liikkumistutkimuksissa ja osa uusia. Kokonaishinta riippuu myös siitä, käytetäänkö uutta tekniikkaa tavallisen tutkimuksen lisänä vai korvikkeena. Vertailutiedon kerääminen kahdella eri menetelmällä (esimerkiksi saman tiedon kerääminen sekä GPS-laitteella että perinteisellä matkapäiväkirjalla tai puhelinhaastattelulla) lisää kustannuksia, kun on maksettava molempien järjestelmien kustannukset.

Henkilökohtaiset haastattelut ja puhelinhaastattelut ovat työvoimavaltaisia, ja niitä käytettäessä työvoimakustannukset muodostavat usein merkittävän osa kokonaiskustannuksista. GPS:n ja muiden automaatiota lisäävien menetelmien käyttö voi siis vähentää kustannuksia; Stopherin ym. (2008) mukaan automatisointi voi säästää aikaa manuaaliseen tiedonkäsittelyyn verrattuna jopa 75–80 %. Uusiin tekniikoihin liittyy kuitenkin alussa tarvittavien ohjelmistojen ja muiden järjestelmien ostamisesta tai kehitystyöstä aiheutuvia suuria kertainvestointeja, jotka maksavat itsensä takaisin vasta pidemmän ajan kuluessa.

GPS:n käytöstä aiheutuvat suurimmat kustannuserät ovat laitteiden hankintakustannukset, tiedonsiirtokustannukset, käytöstä aiheutuvat kustannukset ja tiedonkäsittelykustannukset. Hankintakustannukset riippuvat laitteiden hinnasta ja määrästä. Tiedonsiirtokustannuksia aiheutuu, jos tiedot siirretään langattomasti laitteista tietokantaan. Kustannukset riippuvat tiedonsiirtoyhteyden tyypistä ja nopeudesta, sekä operaattorista riippuen yleensä käytetystä ajasta tai siirretystä tietomäärästä. Joitakin tiedonsiirtohintoja on esitetty liitteessä 2. (Bonsall ym. 2006, paitsi liite.)

Käytöstä aiheutuvia kustannuksia taas ovat esimerkiksi laitteiden käsittelystä (toimitus osallistujille ja takaisin, tietojen lataaminen) ja osallistujien kouluttamisesta ja

neuvomisesta koituvat kustannukset, jotka riippuvat muun muassa posti- tai muista lähetyksmaksuista ja työvoimakustannuksista sekä käytetystä ajasta. Tiedonkäsittelykustannukset puolestaan koostuvat työvoimakustannuksista ja suuren tietomäärän varastointiin ja prosessointiin liittyvistä kustannuksista. Ne riippuvat GPS-osion tarkoituksesta, käytettyjen GPS-laitteiden laadusta, siitä miten fyysinen aineiston varastointi ja prosessointi hoidetaan sekä tarvittavista ohjelmistoista ja tietokannoista (mukaan lukien mahdolliset valmiit paikkatietoaineistot, joiden käytöstä voidaan joutua maksamaan). (Bonsall ym. 2006.)

GPS-laitteiden hinnat vaihtelevat noin 50 €:n ja 500 €:n välillä riippuen siitä, millaiseen laitteeseen GPS-siru on liitetty (Bonsall ym. 2006, appendix A: kattava vertailu raportin kirjoitushetkellä saatavilla olleista erilaisista laitteista (mm. matkapuhelimia ja kämmentietokoneita), hinnat noin 50–500 £). Halvimmat laitteet ovat usein mahdollisimman yksinkertaisia GPS-vastaanottimia; osassa on sisäänrakennettuna tallennuskapasiteettia, osa taas pitää liittää erilliseen tallennuslaitteeseen. Halvempien GPS-laitteiden hintoja ja ominaisuuksia tämän työn kirjoitushetkellä on koottu liitteeseen 1.

Apua henkilöliikennetutkimuksiin muiden alojen sovelluksista?

Paikannussovellusten määrä on loputon, ja lisää sovelluksia eri aloille ja niin yrityksille, tutkimuslaitoksille kuin kuluttajillekin kehitetään jatkuvasti. Lyhyt katsaus joihinkin sovelluksiin on paikallaan, sillä muilta aloilta voidaan saada uusia ideoita liikkumistutkimustenkin tekemiseen ja samoja tai samantapaisia laitteita voidaan hyödyntää eri aloilla. Erityyppisten tutkimusten yhdistämisestä voidaan saada myös synergiaetuja, olettaen että lisätutkimukset voidaan suorittaa aiheuttamatta kohtuutonta lisävaivaa tutkittaville. Seuraava lista ei ole täydellinen, vaan katsaus tarjolla oleviin mahdollisuuksiin.

Liikennetutkimusten lisäksi paikannusta on käytetty myös monilla muilla aloilla apuna ihmisten liikkumisen tutkimiseen. Liikkumista tutkittaessa GPS-laite tyypillisesti annetaan tutkimukseen osallistuvalla joksikin aikaa, ja kaikki hänen liikkumisensa tuona aikana tallennetaan, vastaavasti kuin henkilöliikennetutkimuksissa. Esimerkkejä ovat ihmisten liikkumisen muutosten tutkiminen elämäntilanteen muuttuessa (Flamm ym. 2007), terveyteen liittyvät selvitykset (esimerkiksi Bonsall ym. 2006) sekä lasten liikkumisen ja toiminnan tutkiminen (Kitazawa 2006). Kaupallisemman esimerkin tarjoaa mainosten ohi kävelevien ihmisten lukumäärän selvittäminen paikannuslaitteiden avulla (Pasquier ym. 2008). Etenkin terveyteen ja fyysiseen aktiivisuuteen liittyvissä tutkimuksissa käytetään usein samalla myös erilaisia fyysistä aktiivisuutta mittaavia laitteita, kuten kiihtyvyysantureita (mm. Bonsall ym. 2006, Wolf ja Lee 2008).

Liikennetutkimusten alalla GPS-paikannusta voidaan hyödyntää paitsi henkilöliikennetutkimuksissa myös esimerkiksi matka-aikatutkimuksissa ja nopeusmittauksissa, ja soveltaen myös muun muassa liikenneturvallisuus- ja käyttäytymistutkimuksissa. Kööpenhaminan tiemaksukokeilussa on kerätty GPS-aineistoa, ja satelliittipaikannus onkin yksi varteenotettava vaihtoehto tiemaksujen keräämiseen. (Bonsall ym. 2006.)

GPS-pohjaisesta tiemaksujen keräämisestä kertyisi laaja tietokanta, jota kannattaisi hyödyntää liikenne-ennustekäyttöön. Sitä odotellessa myös muista syistä ajoneuvoihin ja muuhun omaisuuteen kiinnitetyistä GPS-paikantimista syntyy jo nyt laajoja tietokantoja. Niistä saattaisi olla niin paljon hyötyä, että niiden saamisesta tutkimuskäyttöön kannattaisi maksaa. Esimerkiksi ajettuun matkaan perustuvat vakuutukset ja yritysten autojen reaaliaikainen seuranta ovat jo todellisuutta. Tällaisista tietokannoista on mahdollista analysoida myös muutoksia pidemmältä ajalta. (Bonsall ym. 2006.)

Ajoneuvotietokannoissa on tietysti se ongelma, että niistä saadaan tietoja ainoastaan ajoneuvomatkoista. Ihmisten seuraaminen taas onnistuisi tällä hetkellä lähinnä matkapuhelimia paikantamalla, mutta matkapuhelintakaan ei ole kaikilla ihmisillä. Tietojen prosessointiin ja lisätietojen, kuten sosioekonomisten ja demograafisten tietojen, keräämiseen tarvitaan yhteistyötä tietokannan omistajan tai puhelinoperaattorin kanssa, ja tietosuojakysymykset on otettava huomioon. (Bonsall ym. 2006.)

Yksityisyyden suoja

Paikannuksen on usein epäilty olevan sikäli huono tutkimusmenetelmä, että ihmiset ovat tarkkoja yksityisyydestään eivätkä halua itseään valvottavan tai seurattavan. Tähänastiset tutkimustulokset eivät kuitenkaan täysin tue tätä käsitystä. USA:ssa henkilöliikennetutkimuksissa on ollut GPS-osuuksia 1990-luvun puolivälistä asti, eikä ole todisteita, että GPS-osuuteen suostuvien määrä olisi pienentynyt vaikka tieto GPS-tekniikasta on lisääntynyt vuosien kuluessa (Bonsall ym. 2006). Myöskään Australiassa tehdyissä tutkimuksissa yksityisyyssuojat eivät ole nousseet esiin; tosin kyseessä voi vielä olla uutuudenviehätys laitteita kohtaan (Stopher 2004). Osallistumishalukkuus GPS-tutkimuksiin on kuitenkin pienempää kuin perinteisiin matkapäiväkirja-tutkimuksiin (ks. kohta Tutkimusten vastausprosentit, s.37), ja olisikin syytä tutkia, kuinka suuri osa tästä johtuu huolesta yksityisyyden suojasta.

Oletettua pienemmät yksityisyyden suojaan liittyvät ongelmat saattavat olla merkki myös siitä, että kun joku ylittää suostuu mukaan tutkimukseen ja raportoimaan matkansa, hänen yksityisyytensä kannalta ei enää ole niin paljon väliä, millä tavalla tiedot kerätään. On kuitenkin huomattava, että tähän mennessä tutkimukset ovat perustuneet täysin vapaaehtoisuuteen, eli ihmisiä ei ole paikannettu ilman heiltä etukäteen saatua suostumusta. Tällöin paikannuksen mahdollistavan laitteen, oli se sitten GPS-paikannin tai matkapuhelin, voi aina sulkea tai jättää kotiin, jos jostain syystä ei halua matkaansa muiden tietoon (esim. Stopher 2004). Lisäksi Stopher (2008) mainitsee, että ihmiset eivät ehkä ole niin kovin huolissaan yksityisyydestään, kun paikannuksen ja matkoista esitettävien lisäkysymysten välillä kuluu jonkin aikaa, mutta jos joku jo matkan aikana seuraisi ihmisten liikkeitä reaaliajassa ja kyselisi matkaan liittyviä lisätietoja, huolestuneisuus ja vastustus voisi olla paljon suurempaa.

Vaikka edellä todettiin, että tähän mennessä tehdyt GPS-tutkimukset eivät ole herättäneet osallistujissa (ainakaan niissä, jotka ovat suostuneet osallistumaan) juurikaan huolta yksityisyyden suojasta, tilanne on erilainen laajoihin tietokantoihin automaattisesti kertyvien tietojen osalta. Tällöin lupa ihmisten seuraamiseen hankitaan sopimuksella pitkäksi aikaa, esimerkiksi matkapuhelinliittymää hankittaessa tai auton leasingisopimusta tehtäessä. Näin ollaan jo paljon lähempänä ympärivuorokautista valvontaa ja isovelji valvoo -uhkakuvia kuin päivän tai viikon kestävässä tutkimuksessa. Suomessa matkapuhelinverkkoja rakennettaessa ei vielä esiintynyt laajaa vastustusta niihin sisältyvää paikannusmahdollisuutta kohtaan (ehkä myös siksi, että potentiaaliset hyödyt yksittäiselle kuluttajalle olivat niin suuret, ja siksi että tietoja ei saa luovuttaa vapaasti). Sen sijaan joukkoliikennematkojen kartoittaminen elektronisten matkakorttien tietoja tallentamalla ja tiemaksujen kerääminen kaikkiin autoihin asennettavien satelliittipaikantimien avulla ovat herättäneet vastustusta ja pelkoa yksilöiden seuraamisesta siitä huolimatta, että järjestelmät voidaan haluttaessa rakentaa niin, että tallennettuja tietoja ei voida jälkikäteen liittää yksittäiseen ihmiseen.

Pääkaupunkiseudulla vuonna 2008 tehdyssä ajoneuvoliikenteen määräpaikka-tutkimuksessa, jossa kuvattiin autojen rekisteritunnuksia ja lähetettiin niiden perusteella postitse kyselylomakkeita, saatiin hieman kokemusta ihmisten reaktioista uusiin tutkimusmenetelmiin. Kokemusten perusteella voidaan sanoa, että tutkimusmenetelmän

laillisuuden selvittäminen ja tietojen keräämisen tarkoituksen välittäminen osallistujille ovat hyvin tärkeitä. (YTV 2009.)

YTV:n tutkimuksessa moni vastaaja jätti tarkat osoitteet ilmoittamatta, vaikka merkitsikin lähtö- ja määräalueet rastilla annettuun listaan, eli on mahdollista, että he eivät ole halunneet ilmoittaa tarkkaa osoitetta (YTV 2009). Tämä saattaa olla merkki siitä, että paikantamiseen perustuvia tutkimuksia vastustetaan, jos koetaan että niillä saadaan kerättyä liian tarkkaa tietoa. Toisaalta hyvä puoli voi olla, että jos ihmiset saadaan suostuteltua osallistumaan, tiedon tarkkuus paranee.

Yhteenvetona voitaneen sanoa, että vaikka yksityisyyden suojaan liittyviä ongelmia ei vielä ole juurikaan esiintynyt, se ei takaa, etteikö niitä esiintyisi tulevaisuudessa menetelmien kehittyessä ja ihmisten tietämyksen lisääntyessä. Vaarana on, että paikannusta apuna käyttäviin tutkimuksiin liitetään negatiivisia mielikuvia, kuten on jo tapahtunut matkakorteille ja tiemaksuille, ja tutkimusten vastausprosentit putoavat entisestään. Jotta osallistujien rekrytointi mukaan laajoihin tutkimuksiin olisi mahdollista tulevaisuudessakin, rekrytointivaiheessa on rekrytoitavat kyettävä vakuuttamaan siitä, että tutkimus toteutetaan niin, ettei siihen osallistuminen vaarana yksityisyyttä. Yksityisyyden suojaan tulee kiinnittää huomiota paitsi tutkimuksia suunniteltaessa myös etenkin niistä tiedotettaessa.

Tulevaisuuden näkymiä

Jos GPS-tutkimusten todetaan pärjäävän vertailussa perinteisille menetelmille, tulee jossain vaiheessa esille menetelmien tuotteistaminen. Stopherin (2008) näkemys on, että tällä hetkellä ohjelmistoja GPS-aineiston tulkitsemiseen kehitetään monissa paikoissa, mutta jotta GPS-laitteiden käyttö voisi levitä laajemmalle, tarvitaan jonkinlaisia standardoituja ohjelmistoja päättelysääntöineen, tekoälyineen ja sumean logiikan sovelluksineen. Nykyisin ongelmana on myös, että eri maissa tehty liikku- mistutkimukset eivät ole vertailukelpoisia keskenään (Marchal ym. 2008). GPS:ää käyttävä ajankäyttötutkimus täydennettynä esimerkiksi ajoneuvoihin liittyvillä kysymyksillä ja siitä standardoiduilla ohjelmistoilla tehty päätelmät voisivat olla yksi tie kohti vertailukelpoisia tuloksia esimerkiksi EU:ssa.

Tiedonkeruun automatisointi ei kuitenkaan lupaa pelkkää hyvää. GPS:llä ja muilla automatisoiduilla menetelmillä voidaan helposti kerätä hyvinkin suuria määriä aineistoa, jonka prosessoin vie paljon resursseja (Stopher 2008). Tämänkin takia on tarkasti mietittävä etukäteen, millaista aineistoa ja kuinka paljon on kannattavaa kerätä, etenkin jos samalla pyritään standardoimaan tiedonkeruuta useissa maissa kerralla.

Etukäteisarviot uuden tekniikan mahdollisuuksista ovat yleensä positiivisia. Myös kriittinen tarkastelu ja vertailu pitäisi muistaa (mm. Stopher 2008, Bricka 2008). Nyt GPS-tekniikka on edennyt siihen pisteeseen, ettei enää tarvitse vain kokeilla, vaan vertailu perinteisin menetelmin kerättyyn tietoon on mahdollista. Onko elektronisilla päiväkirjoilla oikeasti saatu parempaa ja tarkempaa tietoa? Muistavatko ihmiset oikeasti ottaa GPS-laitteen mukaan jokaiselle matkalle vai eivät?

3.2.2 Katsaus toteutettuihin GPS-avusteisiin liikkumistutkimuksiin

Yleistä

Seuraavaksi esitellään tarkemmin toteutettuja tutkimuksia. Perustietoja neljästä eri tutkimuksesta on esitetty taulukossa 3. Taulukon tutkimuksissa osallistujamäärät ovat melko suuria, mukana ei siis ole pieniä tutkimuksia, joiden tarkoituksena on ollut ainoastaan kerätä aineistoa algoritmien kehittämistä varten. Lisäksi tarkastellaan

tarkemmin muun muassa GPS:n hyväksyttävyyttä, tutkimusten vastausprosentteja, otosta, vastaajille aiheutuvaa vaivaa, käytännön toteutusta ja tutkimusten ongelmia.

Taulukko 3. Muutamien GPS-tutkimusten ominaisuuksia esimerkinomaisesti esiteltynä

Alue (tilaaja, toteuttaja tms.)	Hollanti (3 kaupunkia: Amersfoot, Veenendaal ja Zeewolfe)	Ranska	Ohio, USA (tilaajana The Ohio Department of Transportation, toteuttajana Abt SRBI)	Sveitsi (3 kaupunkia: Zürich, Winterthur ja Geneve)
Tutkimus	Osa suurempaa, asuinpaikan valintaan liittyvää tutkimusta	French National Travel Survey (valtakunnallinen henkilöliikennetutkimus), GPS-osuus	Täysin GPS-pohjainen liikkumistutkimus (household travel survey)	GPS-aineisto on peräisin yksityisen yrityksen tutkimuksesta, jossa selvitettiin, minkä mainostaulujen ohi ihmiset kulkivat
Tallennusväli	6 s	10 s	-	-
Milloin toteutettu	Vuoden 2007 alkupuoliskolla	2007-2008	Tulossa (lehdistötiedote aiheesta joulukuussa 2008)	-
Osallistujia GPS-osiossa	1200 osallistujaa, joista 1104 suoritti tutkimuksen loppuun asti	800 henkilöä	4000 kotitaloutta	4882
GPS-osion pituus	1 viikko	(useita päiviä)	Vähintään 3 päivää	Keskimäärin 6,65 päivää
Validointi	Osallistujat korjasivat virheelliset ja puuttuvat tiedot jälkikäteen internetissä; lisäksi matkojen ominaisuuksia verrattiin Hollannin vuoden 2006 henkilöliikennetutkimuksen tuloksiin.	Kaikkien osallistujien kanssa käydään läpi yhden päivän matkat, joista kysytään matkan tarkoitus, mukana olleiden henkilöiden määrä ja kulkutavat.	1500:a osallistujaa pyydetään jälkikäteen postitse tai internetissä korjaamaan tietoja.	Koska tietoa osallistujista ja todellisista kulkutavoista ei ollut, mutta aineistoa oli paljon, validointi tapahtui vertaamalla tuloksia ko. kaupunkien otokseen vuoden 2005 liikkumistutkimuksessa (Swiss Micro-census on Travel Behaviour 2005).
Lähde	Bohte ja Maat 2008	Marchal ym. 2008	Abt Associates Inc. 2008	Schuessler ja Axhausen 2008

Otos ja sen rajaaminen

Useissa GPS-tekniikkaa ja sen mahdollisuuksia kartoittavissa pilottitutkimuksissa on käytetty vapaaehtoisia osallistujia, joita on rekrytoitu erilaisin keinoin, eikä ole välitetty siitä, että saataisiin koko väestöä edustava otos. Tämä on ymmärrettävää, kun tarkoituksena on todistaa tekniikan toimiminen, kerätä aineistoa algoritmien ja menetelmien kehittämiseen tai vertailla saatuja tietoja perinteisellä matkapäiväkirjalla kerättyihin tietoihin. Siinä vaiheessa, kun halutaan siirtyä laajamittaisiin henkilöliikennetutkimuksiin, on kuitenkin otettava huomioon myös otokseen liittyvät asiat. Seuraavassa käsitellään vastaajien ikää, otoksen kokoa ja työkseen liikkuvia ihmisiä. Kohdassa GPS:n hyväksyttävyys eri ihmisryhmissä käsitellään lyhyesti myös vähemmistöjä.

Otoksen kattavuuteen liittyy paitsi kysymys siitä, miten paljon GPS-tutkimuksiin osallistuvien ominaisuuksien jakaumat poikkeavat koko väestön ominaisuuksien jakaumista, myös se, miten otokset poikkeavat perinteisillä menetelmillä tehtävien tutkimusten otoksista (Bricka 2008). Pitkän aikavälin kehityksen arvioimiseksi on olennaista, että GPS-avusteisista tutkimuksista saadaan aineistoa, joka on vertailukelpoista aiemmin tehtyjen tutkimusten kanssa (Marchal ym. 2008).

Otoksen kattavuuteen ja vertailtavuuteen liittyy myös se, kuinka nuorille laite voidaan antaa, niin eettisistä syistä kuin turvallisuuden ja laitteiden tarkoituksenmukaisen toiminnankin kannalta. Ajoneuvoon asennettavien laitteiden kohdalla ikäkysymystä ei ole tarvinnut miettiä, mutta kannettavien laitteiden myötä myös ikään täytyy ottaa kantaa.

Eri tutkimuksissa alaikärajat GPS-tutkimukseen osallistumiselle vaihtelevat paljonkin. Ranskassa valtakunnallisen henkilöliikennetutkimuksen ensimmäiseen, koeluontoiseen GPS-osioon otetaan vain yli 17-vuotiaita (Marchal ym. 2008). Lee ja Wolf (2008) taas raportoivat useista eri tutkimuksista, joissa ikärajat ovat olleet erilaisia: Chicagon alueen liikennetutkimuksessa mukana oli yli 15-vuotiaita, Massachusettsissa terveystä ja aktiviteettia kartoittavassa tutkimuksessa taas yli 17-vuotiaita. Ohiossa puolestaan on alkamassa liikkumistutkimus, jossa GPS-laitteen saavat yli 12-vuotiaat osallistujat (Abt Associates Inc. 2008).

Useissa Australiassa tehdyissä tutkimuksissa GPS-laitteita on eettisistä syistä annettu vain yli 14-vuotiaille. Pelkona on lisäksi, että sitä nuoremmat eivät välttämättä pitäisi laitetta mukana koko aikaa tai voisivat lainata laitetta kavereilleen, tai laitteet voisivat vahingoittua. Olisi syytä tutkia, kuinka nuorille laite kannattaa antaa. (Stopher 2008.)

Toisaalta on tehty tutkimuksia, joissa nimenomaan on tutkittu lapsia, ja niissä GPS-laitteita on annettu jopa 7-vuotiaille. FACES tutki astmaattisia 7–18-vuotiaita lapsia Californiassa (Wolf ja Lee 2008). CAPABLE taas tutki lasten aktiviteetteja ja lähiympäristön käyttöä Isossa-Britanniassa ja sen osallistujat olivat 8–11-vuotiaita (Kitazawa 2006). Kitazawa (2006) raportoi, että laitteiden ja latureiden käytössä oli ongelmia – useasta laitteesta irtosi hihna, jolla laitetta pidettiin ranteessa, ja muutamasta laitteesta irtosi nappuloita – mutta ei ota kantaa siihen, kuinka suuri osa ongelmista johtui siitä, että laitteita käyttivät lapset. Stopher (2008) analysoi, että CAPABLE-tutkimus onnistui luultavasti osittain siksi, että siinä useammalla saman luokan oppilaalla oli GPS-laite käytössä yhtä aikaa, mikä vähensi luokkatoverien kiinnostusta laitteella leikkimiseen.

Suomessa esimerkiksi valtakunnallisessa henkilöliikennetutkimuksessa mukana ovat kuusivuotiaat ja sitä vanhemmat (Henkilöliikennetutkimus 2004–2005). Jos oletetaan, että GPS-laite annetaan esimerkiksi yli 14-vuotiailla, on tiedot sitä nuorempien matkoista kerättävä muilla menetelmillä vertailukelpoisuuden säilyttämiseksi. Esimerkiksi Ohiossa alkavassa tutkimuksessa tiedot alle 13-vuotiaiden matkoista kerätään paperisella matkapäiväkirjalla (Abt Associates Inc. 2008). Näyttää myös siltä, että mitä vanhempi ihminen on kyseessä, sitä epätodennäköisempää on, että hän suostuu käyttämään GPS-laitetta (Marchal ym. 2008), joten myös vanhojen ihmisten matkat saattavat jäädä vähäisiksi GPS-aineistossa. Toisaalta voidaan myös tutkia, miten liikennemalleihin ja kerätyn aineiston muuhun käyttöön vaikuttaa, jos aineistosta puuttuvat alle tietyin ikäisten lasten matkat (Stopher 2008).

Lisäksi on huomioitava otoskoko. Otoksoon pienentäminen verrattuna perinteisiin tutkimuksiin on mahdollista, jos aineistoa kerätään useammalta päivältä. Tällöin ongelmaksi voi kuitenkin muodostua se, että matkojen määränpäissä ei ole riittävästi vaihtelua matkojen suuntautumisen mallintamista varten. Aineiston keräämisellä useamman päivän ajalta voidaan kuitenkin saavuttaa muita hyötyjä, kuten tietoa

matkojen toistuvuudesta ja liikenteen rytmin vaihtelusta toistuvien matkojen aikana, ja lisäksi toistuvien matkojen avulla voidaan täydentää GPS-aineistosta mahdollisesti puuttuvia kohtia. (Stopher 2008.)

Työkseen liikkuvat ihmiset (esimerkiksi posteljoonit, bussinkuljettajat) ovat vielä yksi haaste. Usein aineistoon ei haluta matkoja, joita he tekevät työtään tehdessään (Henkilöliikennetutkimus 2004-2005). Perinteisissä haastatteluissa on helppo ottaa huomioon vain matkat, jotka eivät ole työntekoa. Jos sen sijaan käytetään GPS-laitetta, joka on päällä koko päivän, tai menee itsestään päälle liikettä havaitessaan, tulevat aineistoon mukaan kaikki matkat, ja työstä aiheutuvat matkat on jotenkin poistettava jälkikäteen. Jos aineistoa tutkimuspäivien jälkeen näytetään osallistujille ja pyydetään korjaamaan virheet (*prompted recall*), voidaan ylimääräiset matkat poistaa silloin, mutta tästä aiheutuu vastaajille ylimääräistä vaivaa. Toinen vaihtoehto on luottaa siihen, että osallistujat muistavat pitää laitteen pois päältä työpäivän aikana.

Tutkimusten vastausprosentit

Perinteisessä posti- tai haastattelututkimuksessa alkuperäisestä otoksesta putoaa ihmisiä pois kahdessa vaiheessa: ensin ne, jotka eivät suostu osallistumaan tutkimukseen ollenkaan, ja sitten ne, joiden antamista tiedoista ei saada hyödyllisiä tuloksia. Samaa lähestymistapaa voidaan käyttää myös GPS-tutkimuksissa (esimerkiksi Stopher 2008, Bricka 2008), jolloin voidaan laskea kuinka monta prosenttia suostuu osallistumaan GPS-tutkimukseen, ja kuinka monta prosenttia lopulta palauttaa hyödyllistä aineistoa. Wolf ja Lee (2008) kuitenkin osoittavat, että GPS-tutkimus jakaantuu huomattavan moneen vaiheeseen, joissa kaikissa vastaaja voi pudota otoksesta.

Wolf ja Lee (2008) erottavat eräistä tekemistään tutkimuksista seuraavat 6 vaihetta (Wolfia ja Leetä mukaillen):

- 1) *Rekrytointi*: Ensimmäinen rekrytointiyritys määrittää, suostuuko rekrytoitava osallistumaan GPS-osuuteen.
- 2) *Osoitteen varmistus*: Jos rekrytoitava antaa suostumuksensa kohdassa yksi, häneltä varmistetaan osoite, johon GPS-laite toimitetaan. Tässä vaiheessa joitain osallistujia ei enää tavoiteta, ja jotkut kieltäytyvät sittenkin osallistumasta.
- 3) *Laitteen toimitus*: Jos osallistuja antaa osoitteensa kohdassa 2, laite viedään tai lähetetään hänelle. Jotkut kieltäytyvät vielä tässä vaiheessa tai eivät ole kotona sovittuna toimitusaikana. Osa laitteista voi myös hävitä postissa.
- 4) *Laitteen käyttö*: Kun laite on saatu toimitettua, osallistujan pitää muistaa ja haluta käyttää laitetta sovittuina päivinä. Osa ei eri syistä tee niin, jolloin aineistoa ei saada.
- 5) *Laitteen palautus*: Riippumatta siitä, noudetaanko laitteet osallistujien kotoa vai voivatko he lähettää ne takaisin, pieni osa osallistujista ei palauta laitteita.
- 6) *Matkapäiväkirjan täyttäminen*: Tutkimuksen tarkoituksena voi olla, että osallistujat raportoivat esimerkiksi yhden päivän matkat matkapäiväkirjaan tai muulla tavoin, jotta niitä voidaan verrata GPS-aineistoon. Tällöin on kokeen onnistumisen kannalta ratkaisevaa, että näin myös tehdään.

Edellä lueteltuihin GPS-tutkimuksen vaiheisiin liittyy myös muita asioita, jotka voivat vaikuttaa vastausprosentteihin ja tutkimuksen onnistumiseen. Rekrytointivaiheessa voidaan käyttää kahta vaihtoehtoista menetelmää: opt-in tai opt-out. Laitteen käyttöön liittyy paitsi ihmisten muistaminen, myös laitteen toiminta. Voi olla, että laitteen toimimattomuuden takia aineistoa ei saada, vaikka kaikki muu menisi hyvin. Kohta 6

liittyy myös siihen, toteutetaanko GPS-tutkimus perinteisen tutkimuksen lisäksi vai sijasta. Vaikka matkapäiväkirjaa ei tarvitsisikaan täyttää, osallistujista tarvitaan yleensä kuitenkin jotain taustatietoja, joiden puuttuminen rajoittaa GPS-aineiston käyttökelpoisuutta. Lisäksi suhtautuminen tutkimuksiin vaihtelee maittain, eli muualla tehtyjen tutkimusten vastausprosentteista ei kannata suoraan tehdä johtopäätöksiä.

Edellä kerrotut seikat huomioiden on selvää, että vastausprosentin määrittäminen ei ole yksiselitteistä. Lisäksi suurin osa tähän mennessä toteutetuista GPS-tutkimuksista on järjestetty perinteisten henkilöliikennetutkimusten yhteydessä. Tällöin rekrytoitavien on yleensä ensisijaisesti haluttu osallistuvan matkapäiväkirjatutkimuksiin, ja jos he ovat suostuneet siihen, heille on tarjottu mahdollisuutta osallistua lisäksi GPS-tutkimukseen. Tällöin osallistumaan suostuvien määrään vaikuttaa se, että he ovat jo suostuneet osallistumaan matkapäiväkirjatutkimukseen. (Stopher 2008.)

Vastausprosentteja tutkimuksista ei ole dokumentoitu kovinkaan hyvin. Lisäksi ongelmana on, että usein on etukäteen päätetty, kuinka monta ihmistä tai kotitaloutta GPS-osioon halutaan mukaan. Tällöin rekrytointia on jatkettu niin pitkään, että ihmisistä on saatu tarpeeksi, mikä vaikeuttaa entisestään tulkintaa vastausprosentteista. Seuraavassa on kuitenkin joitakin arvioita vastausprosentteista sen perusteella, mitä tutkimuksista on dokumentoitu.

Tiedetään, että kahdessa perinteisen tutkimuksen ohessa tehdyssä GPS-tutkimuksessa (Ranskan henkilöliikennetutkimus FNTS ja Australiassa, Victoriassa tehty validointitutkimus) noin kolmasosa matkapäiväkirjatutkimukseen suostuneista suostui osallistumaan myös GPS-osioon. Näistä noin 75 % suoritti tehtävät loppuun eli palautti käyttökelpoista aineistoa. Tästä ei tietenkään voida päätellä, kuinka moni suostuisi osallistumaan pelkkään GPS-tutkimukseen. Muista tutkimuksista saatujen tulosten perusteella vastausprosentit pelkkään GPS-tutkimukseen ovat kuitenkin samaa luokkaa kuin perinteisten tutkimusten. (Stopher 2008).

Ranskan henkilöliikennetutkimuksen (FNTS) yhteydessä on myös käynyt ilmi, että hyväksymisprosentit varsinaisessa tutkimuksessa poikkeavat tutkimusta edeltäneestä pilotista ja muista pienistä tutkimuksista. Ranskassa kysyttiin, suostuisivatko ihmiset osallistumaan GPS-tutkimukseen, ja vastausvaihtoehdot olivat ”Kyllä, ilman ehtoja”, ”Kyllä, sillä ehdolla että laitteen voi laittaa pois päältä”, ”Kyllä, mutta jollain muulla ehdolla” ja ”Ei”. Varsinaisen henkilöliikennetutkimuksen kuudesta aallosta neljässä ensimmäisessä 67 % vastasi ”Ei”, kun taas FNTS-pilotissa vastaava luku oli 60 %, ja Lillessä toteutetussa pienessä tutkimuksessa ”Ei” vastasi vain 50 %. Merkillepantavaa on myös, että kun FNTS-pilotissa lopulta tarjottiin GPS-laitetta ihmisille, jotka olivat vastanneet ”Kyllä” tai ”Kyllä, sillä ehdolla että laitteen voi laittaa pois päältä”, 19 % kieltäytyi siinä vaiheessa sittenkin osallistumasta. (Marchal ym. 2008.)

Ihmisiä voidaan rekrytoida tutkimukseen kahdella eri tavalla, opt-in ja opt-out. Usein matkapäiväkirjatutkimuksen lisäksi järjestetään GPS-osuus, johon halutaan mukaan osa päiväkirjan täyttävistä. Tällöin opt-in GPS-tutkimuksen osalta tarkoittaa sitä, että matkapäiväkirjatutkimukseen suostuville tarjotaan myös mahdollisuutta osallistua GPS-osioon, ja he voivat sanoa haluavansa mukaan. Opt-out-vaihtoehdossa taas valitaan etukäteen osa otoksesta GPS-osioon. Jos he suostuvat mukaan matkapäiväkirjatutkimukseen, heille kerrotaan, että heidät on valittu myös GPS-osioon. Heidän on siis erikseen kieltäydyttävä, jos he eivät halua osallistua GPS-osioon. (Esimerkiksi Bricka 2008.)

Molempia rekrytointitapoja on käytetty (esimerkiksi Bricka 2008, Wolf ja Lee 2008), mutta niiden keskinäisestä paremmuudesta ei ole tehty paljoa johtopäätöksiä. Bricka (2008) toteaa, että USA:ssa molemmilla tavoilla toteutettujen GPS-tutkimusten vastaajien ominaisuuksien jakaumat poikkesivat perinteisellä tavalla toteutetun

tutkimuksen vastaajien ominaisuuksista. Ainoa selkeä ero opt-in- ja opt-out-menetelmien välillä oli kuitenkin se, että opt-out-menetelmällä mukaan saatiin suurempi osuus suuria kotitalouksia, joissa oli lapsia (Bricka 2008). Rekrytointitavasta riippuvaa eroa vastaajissa on siis syytä tutkia vielä lisää.

Joissain tapauksissa ihmisten osallistumishalukkuutta niin GPS- kuin muihinkin tutkimuksiin voidaan kasvattaa tarjoamalla osallistujille palkintoja tai palkkioita, esimerkiksi rahaa tai tavaraa. Käytäntö ja suhtautuminen tällaisiin kannustimiin vaihtelevat maittäin (Stopher 2008), joten niiden vaikutuksia on jo siksi vaikea tutkia. Stopher (2008) listaa yhden yhdysvaltalaisen GPS-tutkimuksen, jossa käytettiin 50 \$:n palkkiota, sekä useita muun muassa australialaisia tutkimuksia, joissa kannustimia ei käytetty, mutta ei yritä näiden pohjalta tehdä johtopäätöksiä kannustimien vaikutuksesta osallistumishalukkuuteen tai lopulliseen vastausprosenttiin. Wolf ja Lee (2008) raportoivat yhdysvaltalaisista GPS-tutkimuksista, joissa käytettiin 10 \$:n ja 50 \$:n palkkioita, mutta eivät analysoi niiden vaikutuksia muuhun kuin laitteiden palauttamiseen. Suomessa palkkioita ei yleensä ole käytetty, mutta joissain tutkimuksissa osallistuneiden kesken saatetaan arpoa lahjakortteja tai muita palkintoja (esimerkiksi Henkilöliikennetutkimus 2004–2005¹²).

GPS:n hyväksyttävyys eri ihmisryhmissä

Siitä, millaiset ihmiset ja kotitaloudet suostuvat osallistumaan GPS-tutkimuksiin, tiedetään onneksi jo hieman enemmän kuin vastausprosentista, sillä osallistujien ja kieltäytyjien ominaisuuksia on dokumentoitu ja tutkittu useiden tutkimusten yhteydessä. Esimerkiksi Bricka (2008) viittaa Hawkinsin ja Stopherin Australiassa tekemään tutkimukseen sekä vuonna 2004 tehtyyn Kansas Cityn henkilöliikennetutkimukseen, ja Marchal ym. (2008) analysoivat Ranskan valtakunnallisen henkilöliikennetutkimuksen GPS-osiota. Osallistujien ja kieltäytyjien analysoinnin kautta voidaan löytää keinoja lisätä osallistumishalukkuutta (Bricka 2008).

Brickan (2008) mukaan osallistumishalukkuus GPS-tutkimuksiin riippuu kahden tekijän yhteisvaikutuksesta: osallistumishalukkuudesta henkilöliikennetutkimuksiin ylipäättään ja tekijöistä, jotka vaikuttavat uuden teknologian hyväksymiseen. Brickan (2008) tekemän yhteenvedon mukaan perinteisissä henkilöliikennetutkimuksissa todennäköisyys vastata pienenee, jos vastaaja on pienituloinen, nuori (alle 25-vuotias), kuuluu vähemmistöön, asuu kaupunkiseudulla, koulutustaso on matala tai kotitalous on suuri. Edelleen vastaamistodennäköisyyttä pienentää, jos henkilö tekee keskimääräistä vähemmän matkoja tai ei käytä joukkoliikennettä. Tässäkin tulokset hieman vaihtelevat maittäin. (Bricka 2008.)

Mitä tulee uuden teknologian hyväksymiseen, Bricka (2008) on tehnyt markkinointitutkimuksista ja tietotekniikkakirjallisuudesta yhteenvedon, jonka mukaan uuden teknologian hyväksymistodennäköisyyttä lisää se, että on nuori, miespuolinen, korkeasti koulutettu tai suurituloinen. Lisäksi Bricka (2008) viittaa tutkimustuloksiin, joiden mukaan teknologian hyväksymistä edistää se, että teknologia koetaan hyödylliseksi ja on helppokäyttöistä. Näillä kahdella lähestymistavalla saaduista vastaamistodennäköisyyteen vaikuttavista tekijöistä osa kumoaa toisensa ja osa vahvistaa toisiaan. Yhteisvaikutuksena voidaan olettaa, että GPS-tutkimukseen osallistuvat vähiten sellaiset, joilla on vähän koulutusta tai pienet tulot.

¹² <http://www.hlt.fi/index.htm> -> Tutkimuksen kuvaus -> Lomakkeet ja kysymykset -> saatekirje -> <http://www.hlt.fi/tutkimus/saatekirje.pdf>

Osallistumishalukkuutta nimenomaan GPS-tutkimuksiin on Brickan (2008) mukaan tutkittu (vuoteen 2008 mennessä) kolmessa tutkimuksessa. Näiden kolmen tutkimuksen tuloksia ja Ranskan henkilöliikennetutkimuksen alustavia tuloksia on koottu taulukkoon 4. Kaikissa analysoiduissa tutkimuksissa GPS-tutkimus toteutettiin perinteisen matkapäiväkirjatutkimuksen ohella. Tulokset eivät ole yhteismitallisia, koska ne on tehty eri kriteereillä ja analysoidut GPS-tutkimukset ovat olleet erilaisia, mutta yhteenvedo antaa osviittaa siitä, mitä tekijöitä vähintäänkin kannattaa huomioida tutkimusten suunnittelussa ja analysoida tulevista GPS-tutkimuksista.

Taulukko 4. Tekijöitä, jotka pienentävät osallistumistodennäköisyyttä GPS-tutkimukseen. Eri tutkimuksissa tekijöitä on vertailtu eri tavoin, eivätkä ne ole yhteismitallisia, mutta taulukko antaa osviittaa siitä, millaisia henkilökohtaisia ominaisuuksia ja liikkuvuustietoja kannattaa ottaa huomioon tutkimuksia suunniteltaessa ja analysoidessa.

	Australia (Greater Sydney Metropolitan Regions Household Travel Survey) *	Kansas City (Regional Household Travel Survey), Tutkimus 1 **	Kansas City (Regional Household Travel Survey), Tutkimus 2 **	Ranska (French National Travel Survey) ***
Muussa maassa syntyneet	X			
Kotitalous muu kuin pariskunta tai pariskunta jolla lapsia	X			
Matala koulutus	X			X
Pienet tulot	X	X	X	X ****
Ei ajokorttia	X			
Kotitaloudessa 1 tai yli 2 (ei 2) henkeä	X			
Pienempi talouden koko		X		X
Keskimääräistä vähemmän autoja taloudessa		(X *****)	X	X
Vähemmän työssäkäyviä / auto			X	
Vanha			X	X
Vähemmän ajoneuvomatkoja / talouden aikuinen			X	
Taloudessa ei tietokonetta, ei internetyhteyttä, ei matkapuhelinta				X
Liikkumiskyky alentunut (huono terveys, jokin vamma)				X
Edellisenä päivänä tehty matkojen määrä vähäinen ja matkustettuja km vähän				X

*) Lähde Bricka 2008 (yhteenvedo); GPS-laitetta tarjottiin 89 kotitaloudelle, joista 48 osallistui ja 41 kieltäytyi

**) Lähde Bricka 2008 (yhteenvedo); 3049 kotitaloutta osallistui perinteiseen matkapäiväkirjatutkimukseen, 1640 ilmaisi kiinnostuksensa osallistua GPS-osuuteen ja 228 kotitaloutta palautti lopulta sekä perinteisen että GPS-aineiston

****) Lähde Marchal ym. 2008, analysoitu myös muutamia muita ominaisuuksia kuin tässä mainitut

*****) Pienet tulot tai tuloja ei ilmoitettu

*****) määritelmällisesti, koska tutkimuksessa GPS-laite kiinnitettiin autoon, joten autottomat taloudet eivät voineet osallistua

Taulukko vahvistaa oletusta siitä, että pienet tulot ja matala koulutus vähentävät osallistumistodennäköisyyttä GPS-tutkimuksiin. Samoin se, että ihminen on vanha tai kotitalous pieni, lisää kieltäytymistodennäköisyyttä. Ranskan tulokset vahvistavat Brickan hypoteesia siitä, että uutta teknologiaa (tietokone, matkapuhelin) muutenkin käyttävät ovat myönteisempiä GPS-tutkimuksia kohtaan. Lisäksi keskimääräistä vähemmän autoja omistavat ja keskimääräistä vähemmän liikkuvat ovat vähemmän halukkaita osallistumaan. Marchal ym. (2008) kiinnittävätkin huomiota siihen, että pelkässä GPS-tutkimuksessa on riski yliarvioida ihmisten liikkumista, koska paljon liikkuvat osallistuvat tutkimukseen todennäköisemmin.

Edellä analysoitiin siis ihmisten osallistumishalukkuutta GPS-tutkimuksiin, aineistona perinteisten matkapäiväkirjatutkimusten yhteydessä toteutetut GPS-osiot. Bricka (2008) analysoi paitsi sitä, miten GPS-osion osallistujat eroavat koko väestöstä, myös sitä, miten GPS-osion osallistujat eroavat perinteisen matkapäiväkirjatutkimuksen osallistujista. Edellä jo mainitussa Kansasissa tehdyssä tutkimuksessa GPS-osion osallistujissa oli vielä perinteistä tutkimustakin vähemmän vanhoja, pienituloisia ja vähemmistöjen edustajia (Bricka 2008). Toisessa, Washington DC:ssä tehdyn tutkimuksen GPS-osiossa oli myös perinteistä tutkimusta vähemmän pienituloisia ja vähemmistöjen edustajia, mutta iän suhteen tulos oli eri, sillä nuoria vastaajia oli perinteistä enemmän (Bricka 2008).

Suomessa vuosien 2004–2005 valtakunnallisessa henkilöliikennetutkimuksessa naiset vastasivat keskimäärin hieman miehiä todennäköisemmin ja vähiten todennäköisesti vastasivat nuoret miehet kokonaisvastausprosentin ollessa 67 prosenttia (Henkilöliikennetutkimus 2004–2005 (2006)). GPS-tutkimuksia ajatellen olisi hyödyllistä Suomessakin analysoida liikkumistutkimusten vastausprosentteja ainakin niissä ryhmissä, joiden muissa maissa on todettu vastaavan keskimääräistä epätodennäköisemmin.

Osallistumisen aiheuttama kuormitus vastaajille

Edellä on mainittu osallistujille aiheutuvan kuormituksen (engl. *respondent burden*) vähentäminen yhtenä tärkeänä syynä GPS-laitteiden käyttöönottoon henkilöliikennetutkimuksissa; tavoitteena on siis kasvattaa osallistumisprosenttia ja kunnolla loppuun asti suoritettujen tutkimusten osuutta. On kuitenkin huomattava, että esimerkiksi tutkimuksissa, joissa vertailutarkoituksessa sekä täytetään perinteistä matkapäiväkirjaa että käytetään GPS-laitetta, kuormitus itse asiassa kasvaa, koska GPS-laitteen käyttö tulee perinteisen tutkimuksen lisäksi eikä tilalle (Bricka 2008). Jotta kuormitus todella vähenisi, on matkojen yksityiskohdat kerättävä vain GPS-laitteella, ja kysyttävien kysymysten määrä on pyrittävä minimoimaan (Schuessler ja Axhausen 2008). Lisäksi Bricka (2008) huomauttaa, että osallistujille erilaisista tutkimusmenetelmistä aiheutuvaa kuormitusta pitäisi jotenkin arvioida, jotta sen voidaan osoittaa vähenevän.

Tutkimukseen osallistumisen kuormittavuutta on vaikea mitata tarkasti, mutta sitä voidaan arvioida esimerkiksi vastaamiseen kuluneen ajan perusteella tai määrittämällä, kuinka monta eri asiaa vastaajan pitää tehdä tutkimukseen liittyen, ja painottamalla eri asioita sopivilla kertoimilla (Bricka 2008). Erilaiset ihmiset kokevat kuitenkin eri asiat eri tavalla kuormittavina, joten painokertoimien määrittäminen on osittain arvailua. Eri tutkimuksissa sama vaihe (esimerkiksi osallistujien rekrytointi tai GPS-laitteiden jakelu ja palautus) on myös voitu toteuttaa eri tavalla, jolloin sen kuormittavuus ei välttämättä ole sama.

Bohte ja Maat (2008) raportoivat hollantilaisesta tutkimuksesta, jossa osallistujat kantoivat viikon GPS-laitetta mukana ja sitten käyttivät internetsovellusta, jossa he korjasivat tutkijoiden tekemiä päätelmiä matkoista ja lisäsivät puuttuvia tietoja. Lopuksi osallistujia pyydettiin arvioimaan tutkimuksen eri osioiden kuormittavuutta. Tulokset on

toistettu taulukossa 5. GPS-laitteen käyttämistä ja lataamista ei koettu kovin rasittavaksi, mutta internetsovelluksen käytössä monilla oli vaikeuksia. Tätä tutkijat selittävät sillä, että GPS-aineistossa oli paljon aukkoja, joita vastaajat joutuivat täyttämään. Ihmiset, joilla oli vähän kokemusta tietokoneiden käytöstä tai jotka käyttivät vanhaa tietokonetta, raportoivat muita enemmän ongelmista. Lisäksi selvisi, että 40 % osallistujista koki ongelmaksi muistaa ottaa GPS-laitteen mukaan ja neljännes unohti sen ainakin kerran. (Bohte ja Maat 2008.)

Taulukko 5. Vastaajien kokema tutkimuksen eri osien kuormittavuus hollantilaisessa GPS-tutkimuksessa. Kuormitusta arvioitiin kolmesta eri asiasta: GPS-laitteen kuljettaminen mukana, GPS-laitteen lataaminen ja internetsovelluksen käyttäminen. Vastausvaihtoehdot olivat hyvin kuormittavaa, jossain määrin kuormittavaa ja ei kuormittavaa. (Lähde: Bohte ja Maat 2008, Table 3.)

Osallistujien raportoima kuormitus (N=905)			
	Paljon	Jonkin verran	Ei
GPS-laitteen kantamisesta aiheutui kuormitusta	1 %	14 %	85 %
GPS-laitteen lataamisesta aiheutui kuormitusta	5 %	10 %	86 %
Internet-sovelluksesta aiheutui kuormitusta	22 %	45 %	33 %

Laitteisiin liittyviä käytännön seikkoja ja ongelmia

GPS-tutkimuksissa on kiinnitettävä huomiota myös tutkimustilanteessa esiintyviin käytännön ongelmiin, etenkin niihin, jotka johtuvat laitteista ja joita ei siten esiinny perinteisissä tutkimuksissa. Marchal ym. (2008) nostavat esille seuraavat ongelmat:

- GPS-laitteen energiansaanti (akkujen kesto vaihtelee ja ne pitää muistaa ladata)
- GPS-signaalin vastaanotossa esiintyvät ongelmat
- Osallistujaan liittyvät ongelmat, kuten:
 - Osallistuja voi unohtaa ottaa GPS-laitteen mukaan (joidenkin matkojen tai joidenkin päivien ajaksi).
 - Osallistuja voi lainata GPS-laitetta toiselle henkilölle.
 - Osallistuja saattaa haluta ”leikkiä” GPS-laitteella, jolloin se tallentaa enemmän matkoja kuin pitäisi.

GPS-laitteiden toimittamisessa ja palauttamisessa voi myös olla ongelmia. Jos laitteet viedään osallistujille kotiin, on sovittava toimitusaika, ja voi olla että ketään ei silti ole kotona. Jos taas laitteet postitetaan, voi pieni osa laitteista hävitä matkalla. Tutkimuspäivien jälkeen laitteet voidaan noutaa osallistujien luota tai heitä voidaan pyytää lähettämään se takaisin. Noutamista on yleensä käytetty, jotta mahdollisimman moni laite saataisiin takaisin. Eräässä Chicagossa tehdyssä tutkimuksessa, jossa osallistujien piti lähettää laite takaisin, useita laitteita ei ikinä palautettu toistuvista yhteydenotoista huolimatta. Edes tutkimuksesta laitteen palauttamisen jälkeen maksettu 10 \$:n palkkio ei ollut riittävä kannustin laitteen palauttamiseen. (Wolf ja Lee 2008.)

Joka tapauksessa laitteiden toimittamiseen ja palauttamiseen kuluu aikaa, ja aikataulussa on varauduttava viivytyksiin. GPS-laite voi myös olla joillekin osallistujille liian vaikea käyttää, etenkin jos laite lähetetään postissa eikä osallistujille voida näyttää, miten sitä käytetään. Laite voi myös lakata toimimasta kesken tutkimuspäivän, jolloin menetetään pahimmassa tapauksessa myös jo siihen mennessä kerätty aineisto.

Dokumentointi ja ohjeistus

GPS-laitteiden käyttö liikkumistutkimuksissa on vielä niin uusia asia, että suuri osa tutkimuksista on ollut pieniä ja epästandardeja. Useissa julkaistuissakin tutkimuksissa

otoksen poimintatapa, vastausprosentti ja millaiset ihmiset ovat vastanneet sekä käytetyt kannustimet on dokumentoitu huonosti tai ei ollenkaan (Bricka 2008).

Lisäksi olisi hyödyllistä tietää, miten tutkimus esiteltiin osallistujille, millaiset ohjeet heille annettiin, ja miten neuvottiin toimimaan, jos laite ei toimi kunnolla. Tällaiset yksityiskohdat voivat auttaa selvittämään, mitkä tekijät kasvattavat ja vähentävät GPS-tutkimuksiin osallistumista ja tehtävän suorittamista kunnolla. (Bricka 2008.)

3.3 Algoritmit GPS-aineistolle

3.3.1 GPS-aineiston automaattinen käsittely

Yleistä algoritmeista ja niiden hyödyntämisestä

Pieniä määriä GPS-aineistoa voidaan käsitellä käsin. Näin voidaan tehdä, jos esimerkiksi tutkitaan pientä määrää ihmisiä ja aineistosta pitää selvittää jokin tietty asia. Laajemmissa tutkimuksissa, kuten henkilöliikennetutkimuksissa, osallistujia voi kuitenkin olla kymmeniä tuhansia, GPS-aineistoa kertyy valtavia määriä ja aineistosta halutaan selvittää monia asioita. Aineiston käsittely käsin kestäisi liian kauan ja maksaisi liikaa, joten käsittelyä on automatisoitava. Automatisointi tarkoittaa käytännössä tietokoneohjelmia ja sovelluksia, joissa aineistoon sovelletaan erilaisia algoritmeja.

GPS-aineiston muuttamiseksi hyödylliseen muotoon on kehitetty muun muassa algoritmeja, joilla voidaan arvioida aineiston järkevyyttä ja poistaa virheelliset pisteet, pilkkoa aineisto matkoihin ja eri kulkutavoilla tehtyihin matkan osiin sekä päätellä matkan kulkutapa, reittikadut ja tarkoitus. Vaikka eri tarkoituksiin tehtyjä algoritmeja käsitellään seuraavassa erillisinä, on yleensä tarkoituksenmukaista yhdistää useampia algoritmeja kokonaisuudeksi. Näin on ensinnäkin siksi, että lähtökohtana on yleensä koordinaattijonosta koostuva GPS-aineisto, joka on siivottava virheellisistä pisteistä ja jaoteltava pienemmiksi osiksi ennen kuin sille voidaan tehdä mitään muuta. Lisäksi ainakin kulkutavan, reitin ja matkan tarkoituksen päättely voivat tukea toisiaan.

Ainoastaan kulkutavanpäättelyalgoritmeja käydään tässä yksityiskohtaisesti läpi. Muista algoritmeista käydään läpi joitakin yleisiä asioita, kuten mitä niillä tarkemmin ottaen tehdään, mitä lähtötietoja algoritmeissa voidaan käyttää GPS-aineiston lisäksi, mitä ongelmia päättelyyn liittyy ja miten hyvin algoritmit on saatu toimimaan.

Osa kirjallisuudessa esitellyistä algoritmeista on kehitetty sellaisten tutkimusten pohjalta, joissa GPS-laitteet asennettiin autoihin, ja osa sellaisten pohjalta, joissa tutkittiin useampia kulkumuotoja eli osallistujat kantoivat GPS-laitetta mukana kaikilla matkoilla. Autoihin asennettujen GPS-laitteiden aineistoa on helpompi analysoida, mutta siinä on olennaisia puutteita: muut kuin automatkat ovat olennainen osa ihmisten liikkumista, varsinaisia lähtö- ja määräpaikkoja ei saada selville, koska vain auton reitti tallentuu, ja tieto ajoneuvon kuljettajasta kullakin matkalla pitää kysyä erikseen (Schuessler ja Axhausen 2008). Lopullisena tavoitteena on tietysti, että algoritmeilla voitaisiin tutkia kaikkia kulkutapoja, jotta saadaan kokonaiskuva ihmisten liikkumisesta. Siksi seuraavassa käsitellään pääosin sellaisia algoritmeja, joilla on tutkittu muitakin kuin automatkoja.

GPS-aineiston käsittelyssä lähdetään yksinkertaistetusta oletuksesta, että koehenkilön päivä jakautuu matkoihin ja niiden välissä oleviin aktiviteetteihin; aktiviteetti on siis tässä mitä tahansa toimintaa tai oleskelua, jonka aikana ei matkan määritelmän mukaisesti olla siirtymässä paikasta toiseen. Matkan ja aktiviteetin määrittelemisen ei kuitenkaan aina ole helppoa, sillä osa matkoista ja aktiviteeteistä on hyvin lyhyitä

(esimerkiksi pistäytyminen kauppaan matkan varrella). Tämä vaikeuttaa myös aineiston tulkittamista.

Algoritmeja voidaan usein parantaa käyttämällä vastaajilta kysyttyjä taustatietoja esimerkiksi heidän tavallisesti käyttämistään kulkutavoista ja paikoista, joissa he käyvät. Tässä joudutaan kuitenkin pohtimaan, parantavatko taustatiedot algoritmeja niin paljon, että vastaajia kannattaa kuormittaa niiden kysymisellä.

Käytettävissä olevat kulkutavat, niille ominaiset nopeudet ja muut tekijät vaihtelevat eri alueilla. Myös algoritmien apuna käytettävien paikkatietoaineistojen tarkkuus ja saatavuus vaihtelee. Vaikka jokin algoritmi siis toimisikin hyvin tietyllä alueella, sitä ei välttämättä voida sellaisenaan siirtää muille alueille, joilla olosuhteet, käytettävissä olevat kulkutavat ja kerätyt paikkatietoaineistot ovat erilaisia.

Aineiston esikäsittely, järkevyyden tarkastelu ja virheiden poisto

Ennen aineiston analysointia sen esitystapaa joudutaan usein muuttamaan. Päivämäärä ja kellonaika on syytä muuttaa GPS-laitteen tallentamasta ajasta paikalliseksi ajaksi, kesäaika huomioiden. Päivämäärästä voidaan tarvittaessa laskea viikonpäivä. Lisäksi koordinaattien esitystapaa voidaan joutua muuttamaan, ja muitakin muutoksia luultavasti joudutaan tekemään. Suuri osa muutostarpeista riippuu käytetyistä laitteista ja ohjelmistoista. (Stopher ym. 2008.)

GPS-tekniikasta kertovassa kappaleessa 3.1 käsiteltiin erilaisia GPS-aineistoon virheitä aiheuttavia tekijöitä. Virhelähteiden vuoksi GPS-aineiston varsinainen käsittely aloitetaan yleensä arvioimalla, mitkä pisteet aineistossa ovat valideja, eli sellaisia joiden kautta reitti todella on kulkenut, ja mitkä pisteet ovat virheellisiä. Virheellisiksi tuomitut pisteet poistetaan aineistosta. (mm. Stopher ym. 2008.)

Osa GPS-laitteista tallentaa tarkkuudesta kertovia tietoja kuten käytettävissä olevan satelliittien määrän tai satelliittien keskinäisen aseman, jolloin näitä tunnuslukuja tarkastelemalla voidaan arvioida osa pisteistä epäluotettaviksi (esim. Stopher ym. 2008). Ilman niitäkin voidaan kuitenkin selvittää tarkastelemalla esimerkiksi nopeuksia ja kiihtyvyyksiä ja poistamalla ne pisteet, joissa nopeus tai kiihtyvyys on epärealistisen suuri (Schuessler ja Axhausen 2008) tai poikkeaa liiaksi viereisten pisteiden vastaavasta arvosta.

Täysin virheelliset pisteet pyritään siis poistamaan edellä mainituilla suodatus-tekniikoilla. Sen sijaan aineistossa olevia pieniä, satunnaisia virheitä voidaan vähentää erilaisilla tasoitustekniikoilla (engl. *smoothing techniques*), jolloin esimerkiksi ilmakehän häiriöiden ja rakennusten aiheuttamien heijastusten vaikutuksia voidaan pienentää. (Schuessler ja Axhausen 2008.)

Aineistosta poistetaan usein myös pisteet, joista tulkitaan, että laite on pysynyt paikoillaan. Etenkin silloin, jos laite tallentaa koko päivän ajan, tulee tallennettua paljon turhaa aineistoa. Paikoillaan pysyminen voidaan havaita siitä, että nopeus on lähellä nollaa tai useat pisteet kasautuvat lähes samaan kohtaan (Stopher ym. 2008). Se, miten lähelle pysähtymispaikkaa GPS-laitteen tallentamat pisteet kasautuvat, riippuu laitteen tarkkuudesta ja siitä, kuinka hyvin GPS-signaali kuuluu pysähtymispaikkaan. Stopher ym. (2008) käyttivät rajana 30 metriä sillä perusteella, että GPS:n antaman sijainnin keskihajonnaksi saatiin 10 metriä, jolloin käyttämällä kolmea kertaa keskihajontaa etäisyysrajana saatiin 99,8 % todennäköisyydellä selville, milloin laite lähti taas liikkeelle. Kuvassa 8 on esimerkki siitä, millaista aineistoa GPS-laite tallentaa ollessaan paikoillaan sisätiloissa.



Kuva 8. GPS-laitteen tallentama "reitti" kun laite on ollut useita tunteja sisätiloissa paikoillaan valkoisen nuolen osoittamassa kohdassa. Mustalla on merkitty rakennuksen ääriiviivat. Punaiset pallot kuvaavat kohtia, joissa laitteessa olevaa nappulaa on painettu matkan alun ja lopun merkiksi. (Matkan alkua kuvaavat pallot eivät välttämättä ole kovin tarkasti oikeassa kohdassa, koska laite on juuri tuotu ulos rakennuksesta ja paikannustarkkuus on siis siksak-kuvion luokkaa.) Siksak-kuvion koko vaihtelee eri rakennuksissa sen mukaan, miten hyvin GPS-signaali kuuluu rakennuksen sisään.

Aineistosta puuttuvat osuudet

GPS-aineistolle on tyypillistä, että osa matkasta saattaa puuttua aineistosta. Yksi syy on se, että GPS-laitteen käynnistyessä kuluu vähintään muutamasta sekunnista puoleen minuuttiin (joskus useita minuutteja) satelliittien löytämiseen, ja matkan alku puuttuu siis useimmiten aineistosta. Lyhyet matkat voivat puuttua kokonaan, jos laite ei ehdi löytää satelliitteja ollenkaan. (Stopher ym. 2008.)

Lisäksi signaali häviää tunneleissa ja usein myös joukkoliikennevälineissä (metrossa maan alla aina, ja junassakin melko usein), jolloin aineistoon jää aukkoja keskelle matkaa. Korkeiden rakennusten keskellä tai muissa paikoissa, joissa signaalin kuuluvuus on huono (esimerkiksi alikulkutunneleissa, parkkihalleissa ja ostoskeskuksissa) tallentuvien pisteiden tarkkuus voi olla niin huono, että automaattinen prosessi poistaa osan pisteistä.

Puuttuva matkan alku on usein helpointa päätellä siitä, että matkan näennäinen lähtöpaikka ei täsmää edellisen matkan saapumispaikkaan.

Signaalin häviämisen yhteydessä voidaan selvittää aika, jonka signaali on ollut kateissa, matka jonka se on tänä aikana kulkenut ja paikat joissa signaali katosi ja löytyi taas. Niistä voidaan yrittää päätellä, onko henkilö signaalin ollessa kateissa vain jatkanut

matkaa samalla välineellä kuin siihenkin asti, vaihtanut kulkuvälinettä (metroon, junaan tai muuhun välineeseen, jossa signaali usein häviää) tai suorittanut jonkin aktiviteetin.

Puuttuvat osuudet on tärkeää pystyä rekonstruoimaan jollain tavalla, jotta saadaan oikeampi käsitys matkojen pituuksista ja kestoista esimerkiksi liikennesuunnittelua, vertailua ja päästöjen laskentaa varten. Puuttuva reitti voidaan päätellä esimerkiksi lyhyimmän tien etsivällä algoritmilla, jos käytössä on sopiva paikkatietoaineisto ja ohjelma. Puuttuvan osuuden nopeutta voidaan arvioida käyttämällä sitä edeltäneen tai sitä seuraavan matkan osan keskinopeutta. (Stopher ym. 2008.)

Aineiston jakaminen matkoiksi ja matkan osiksi

Jotta GPS-aineistosta voitaisiin päätellä asioita perinteiseen tapaan, se on ensin jaettava matkoiksi (engl. *trip*), joille voidaan määrittää esimerkiksi lähtö- ja määräpaikka, pituus, kesto ja tarkoitus. Matkan ja matkaketjujen määrittelemiseen liittyy samoja haasteita kuin perinteisissä tutkimuksissakin (esimerkiksi miten suhtaudutaan matkan varrella tapahtuviin lyhyisiin pysähdyksiin kuten ihmisten ja tavaroiden noutaminen tai jättäminen, kahvin ostaminen tai kaupassa poikkeaminen, ja riippuuko suhtautuminen siitä, onko pysähdyspaikka reitin varrella vai pitääkö sitä varten poiketa reitiltä), joten määrittelyihin on automatisoinnin yhteydessä syytä kiinnittää huomiota (Stopher 2008). Kulkutavan päättelyä varten matkat on yleensä vielä jaettava matkan osiksi (engl. usein *trip stage*), jotka siis määräytyvät siten, että kulkutavan vaihto erottaa matkan osat toisistaan.

Aineiston pilkkomisessa matkoiksi ja matkan osiksi voidaan käyttää hyväksi pysähdyksiä ja niiden pituuksia. Se ei kuitenkaan ole yksinkertaista, sillä eripituisia pysähdyksiä syntyy paitsi aktiviteeteista, jotka erottavat matkat toisistaan, myös kulkutavanvaihtoista, jotka erottavat matkan osat toisistaan. Pitäisi siis pystyä jollain todennäköisyydellä päättelemään pysähdyksen pituudesta, onko kysymyksessä aktiviteetti vai kulkutavanvaihto. Liian lyhyen aikavälin käyttö johtaa siihen, että pysähdykset muun muassa liikennevaloissa ja bussipysäkeillä tulkitaan virheellisesti matkan päättymiseksi, kun taas liian pitkän aikavälin käytöstä seuraa, että lyhyitä aktiviteetteja kuten matkustajien noutamista tai pikaruuan ostamista ei huomata (Stopher 2008). Aikaraja, jota pidemmät pysähdykset tulkitaan aktiviteeteiksi, vaihtelee. Schuessler ja Axhausen (2008) löysivät kirjallisuudesta arvoja 45 ja 300 sekunnin välillä – yleisin käytetty arvo oli 120 sekuntia – mutta päätyivät itse käyttämään 900 sekuntia.

Marchal ym. (2008) taas ovat selvittäneet vanhojen Ranskan henkilöliikennetutkimusten perusteella, kuinka monta prosenttia aktiviteeteista ja kulkutavanvaihtoista kestää kuinkakin monta minuuttia. Vaihdoista 44 prosentissa kesto on 0 minuuttia eli vaihto tehdään heti. (Marchal ym. 2008.)

Kun käytettävissä on sopivia paikkatietoaineistoja, voidaan myös pysähtymispaikoista yrittää päätellä, onko lyhyessä pysähdyksessä kyseessä liikennevalot, ruuhka, vaihto tai aktiviteetti. Bussipysäkillä tapahtuva pysähdys on todennäköisimmin vaihto, kun taas koulun tai päiväkodin luona tapahtuva pysähdys indikoi, että joku on todennäköisesti otettu kyytiin tai jätetty kyydistä. Tässä on kuitenkin oltava varovainen sikäli, että kyyditys muistuttaa vaihtoa eli tapahtuu usein asemalla tai pysäkillä ja kestää tyypillisesti myös hyvin lyhyen aikaa. Ihmisten ja tavaroiden noutamiseen liittyviä pysähdyksiä voidaan yrittää tunnistaa myös sillä perusteella, että silloin usein käännetään ympäri 180 astetta ja liikutaan ainakin vähän matkaa takaisin tulosuuntaan (Schuessler ja Axhausen 2008 viittaavat tutkimuksiin, joissa näin on tehty).

Vaikka käytössä ei olisikaan valmista paikkatietoa, josta selviää liikennevalojen sijainnit ja ruuhkautuvat alueet, voisi ehkä olla mahdollista erottaa alueita, joilla

toistuvat pysähdykset ovat odotettavissa (esimerkiksi kaupunkien keskustoissa) ja toisaalta alueita, joilla ei pitäisi tulla ylimääräisiä pysähdyksiä (tosin vasemmalle kääntyminen risteyksissä voi tällöin aiheuttaa ongelmia). Tämän luokittelun perusteella voisi yrittää päätellä, onko lyhyt pysähdys aktiviteetti vai ei. Jos taas käytössä on hyvinkin tarkka paikkatietoaineisto teineen ja risteyksineen ja GPS-aineisto on niin tarkkaa, että voidaan sanoa, milloin liikkuja on ollut tiellä ja milloin ei, voitaneen paikkatiedon perusteella päätellä, milloin tieltä on poikettu lyhyen aktiviteetin takia.

Jos käytettävissä on aineistoa saman ihmisen matkoista useamman päivän ajalta, voidaan toistuvia matkoja verrata toisiinsa ja yrittää päätellä, poikkeako joku muista siten, että voidaan olettaa matkan varrella olleen ylimääräinen pysähdys (Stopher 2008).

Joukkoliikenneterminaaleissa, joissa on kauppoja ja muita palveluja, voi olla vaikea päätellä onko kyseessä pelkkä vaihto vai ei. Oikeastaan kyseessä on myös määritelmällinen kysymys sikäli, että jos tarkoituksena on vaihtaa kulkutapaa ja liikkeisiin mennään tappamaan aikaa kulkuvälineen lähtöä odotellessa, ei kyseessä ole aktiviteetti siinä mielessä, että kyseiseen paikkaan olisi tultu aktiviteetin takia. Joissain tapauksissa siis matkojen osien yhdistäminen kokonaisiksi matkoiksi on erillinen ongelma.

Verrattaessa GPS-aineistosta etsittyjen matkojen määriä perinteisissä matkapäiväkirjaita haastattelututkimuksissa tuloksiksi saatuihin matkamääriin käy usein niin, että matkojen määrä GPS-aineistossa on suurempi kuin perinteisissä tutkimuksissa. Tähän on kaksi pääasiallista syytä. Ensinnäkin algoritmeilla pyritään yleensä löytämään kaikki matkojen päätepisteet, jolloin on todennäköistä, että vääriä tulkintoja esiintyy ja matkojen lukumäärä on suurempi kuin pitäisi. Toiseksi perinteisissä tutkimuksissa lyhyet matkat tai pysähdykset matkan aikana unohdetaan usein mainita, eli ilmoitettujen matkojen määrä on pienempi kuin todellisuudessa. On siis todisteita siitä, että perinteisten tutkimusten luvut ovat liian pieniä, mutta toisaalta GPS-aineistosta matkoja erottelevat algoritmit vaativat vielä kehittämistä ennen kuin niidenkään antamiin lukuihin voidaan luottaa. (mm. Bohte ja Maat 2008, Schuessler ja Axhausen 2008.)

Kulkutavanpäätely

Kulkutavanpäätelyalgoritmien lähtökohtana on yleensä nopeus (keskinopeus, maksiminopeus, nopeuden vaihtelut, nopeus jota on käytetty suurin osa matkasta). Koska eri kulkutapojen nopeusalueet ovat päällekkäisiä, on käytettävä tarkentavia metodeja, kuten kiihtyvyydestietoja, erilaisiin ehtoihin perustuvaa päätelyä tai sumeaa logiikkaa. (mm. Schuessler ja Axhausen 2008, Bohte ja Maat 2008, Stopher ym. 2008.)

GPS-aineiston lisäksi voidaan käyttää paikkatietoaineistoja ja selvittää, onko kuljettu reitti riittävän lähellä jonkin joukkoliikennevälineen reittiä, jotta voidaan olettaa että matka on tehty kyseisellä välineellä. Vaikka koko reitti ei olisikaan tiedossa, voidaan jo vertaamalla joukkoliikenteen pysäkkejä GPS-aineiston pysähdyksiin saada tietoa siitä, voisiko kyseessä olla jokin joukkoliikenneväline. Maan alla GPS-signaali häviää, mutta metromatkat voidaan usein tunnistaa siitä, että edellinen matkan osa päättyy metroaseman lähelle ja seuraava alkaa metroaseman läheltä. (Bohte ja Maat 2008.)

Jos käytettävissä on taustatietoja vastaajista, kuten polkupyörän tai auton omistus tai käyttömahdollisuus, tai julkisen liikenteen lipun omistus, voidaan päätellä, että jotkut kulkutavat ovat poissuljettuja tai epätodennäköisiä (esimerkiksi Stopher ym. (2008) käyttävät päätelyssä tietoa polkupyörän käyttömahdollisuudesta kotitaloudessa).

Kahdessa pienessä tutkimuksessa (toisessa 9 vapaaehtoista ja toisessa 60 matkaa) kulkutavantunnistusalgoritmit on saatu toimimaan niin hyvin, että ne tunnistivat oikein noin 90 % tapauksista. Bohte ja Maat kuvaavat suurempaa tutkimusta, jossa aineistona olivat 1104 ihmisen viikon aikana Hollannissa tekemät yli 33 000 matkaa. Siinä osallis-

tujat eivät täyttäneet matkapäiväkirjaa, vaan GPS-aineisto prosessoitiin algoritmeilla, ja sen jälkeen osallistujat pääsivät internet-sovelluksen avulla korjaamaan ja täydentämään tietoja. Kuljutavanpäättely oli mennyt oikein 70 % tapauksista. (Bohte ja Maat 2008).

Stopher ym. (2008) taas raportoivat, että Australiassa kehitetty algoritmi toimii noin 95 %:ssa tapauksista, ja se tunnistaa kävelyn, pyöräilyn, auton, bussin ja junan. Myös epätavallisemmat kuljutavat, O-bahn ja raitiovaunu Adelaidessa, pystytään tunnistamaan, lähinnä niiden reittien perusteella.

Toteutettuja kuljutavanpäättelyalgoritmeja esitellään tarkemmin kappaleessa 3.3.2.

Reitin päättely

Reitin päättelyn tarkoituksena on sovittaa kuljetettu reitti kartalle (engl. *map matching*) eli käytännössä paikkatietoaineistona olevan tieverkon linkkeihin. Vastaavaa tekniikkaa on kehitetty GPS-pohjaisen navigoinnin ja reitinopastuksen tarpeisiin. (Chung ja Shalaby 2005.)

Reitin päättelyyn tarvitaan tarkka multimodaalinen verkko, tai jos sellaista ei ole saatavilla on käytettävä useita verkkoja eri kuljutavoille (Schuessler ja Axhausen 2008). Helpointa reitin päättely lieenee moottoroitujen kulkuneuvojen osalta, koska tieverkko ja joukkoliikenteen reitit ovat yleensä ainakin jollain tarkkuudella olemassa paikkatietoaineistona. Sen sijaan kevyen liikenteen verkkoa ei välttämättä ole olemassa paikkatietona, jolloin reittiä ei voida sovittaa siihen.

Reitin päättely tukee kuljutavan päättelyä useilla tavoilla. Jos kuljetettu reitti näyttää noudattavan joukkoliikennevälineen reittiä, on todennäköistä, että kyseessä on joukko-liikennematka. Varmistukseksi voidaan etsiä todisteita esimerkiksi odottelusta pysäkin lähellä. Jos taas kuljetettu reitti sisältää moottoriteitä tai rampeja, voidaan olettaa että kyseessä ei ole kevyen liikenteen matka. Jos on käytettävissä tietoa katujen yksisuuntaisuudesta, väärään suuntaan yksisuuntaista katua tehty matka voidaan olettaa kevyen liikenteen matkaksi. (Chung ja Shalaby 2005.)

Määräpaikan tyypin ja matkan tarkoituksen päättely

GPS-aineistosta saadaan edellä esitetyllä tavalla matkojen lähtö- ja määräpaikkojen sijainti. Käyttämällä osallistujilta kysyttyjä taustatietoja ja sopivia paikkatietoaineistoja voidaan päätellä lähtö- ja määräpaikkojen tyypit. Päättelysäännöillä, jo tehdyistä tutkimuksista saatavilla todennäköisyysjakaumilla tai oppivilla järjestelmillä voidaan kehittää sääntöjä, joilla päätellään matkan tarkoitus lähtö- ja määränpään tyypin perusteella.

Onnistumisprosentti riippuu muun muassa siitä, kuinka tarkasti määräpaikan tyyppi ja matkan tarkoitus halutaan selvittää, eli kuinka moneen luokkaan ne yritetään luokitella. Käytettävissä olevien paikkatietoaineistojen laatu ja tarkkuus ovat myös oleellisia. Paikkatietoaineistoista voidaan käyttää esimerkiksi maankäyttöluokituksia yhdistettynä tietokantoihin, joissa on erilaisten kohteiden (koulut, ostoskeskukset, terveyskeskukset ja sairaalat ja vastaavat) koordinaatteja (Stopher 2008).

Stopher ym. (2008) ovat päätyneet tulokseen, että kysymällä taustatietoina osallistujien kotiosoite, työpaikan, koulun tai opiskelupaikan osoite sekä kahden useimmin käytetyn ruokakaupan osoitteet voidaan tunnistaa 75 % määräpaikoista ja päätellä 60 % matkojen tarkoituksista. Loput pyritään päättelemään maankäyttötietojen perusteella, myös tietoa vastaajan ammatista voidaan hyödyntää. Matkan tarkoitukselle oli tässä tapauksessa 10 eri vaihtoehtoa (kotiperäinen 8 kpl: työ / koulu / ostos / ruokailu / henkilökohtainen tai lääketieteellinen / sosiaalinen tai virkistys / kyyditys tai nouto / muu; työperäinen; ei koti- eikä työperäinen). (Stopher ym. 2008.)

Matkan tarkoituksen päättelyssä voidaan käyttää apuna myös muita tietoja, esimerkiksi kellonaikaa ja aktiviteetin kestoa (Schuessler ja Axhausen 2008).

3.3.2 Kulkutavanpäättelyalgoritmit

Kulkutavanpäättelyyn on useita eri lähestymistapoja: päättelysääntöjä, todennäköisyyksiä, painokertoimia, sumeaa logiikkaa, oppivia algoritmeja ja näiden yhdistelmiä. Useimmissa tapauksissa ei tietenkään voida saada täysin varmasti selville, mitä kulkutapaa on todellisuudessa käytetty milläkin matkan osalla. Tällöin, jos jokin kulkutapa vaikuttaa riittävän todennäköiseltä, asetetaan se kyseisen matkan osan kulkutavaksi. Joidenkin osien kohdalla voi käydä niin, että mikään kulkutapa ei vaikuta riittävän todennäköiseltä, jolloin kulkutapa voidaan myös jättää määrittämättä.

Seuraavaksi kuvataan neljää kulkutavanpäättelyalgoritmia tarkemmin. Ennen jokaisen algoritmin kuvausta esitetään taulukkona joitakin taustatietoja algoritmista, sen toimivuudesta ja tutkimuksesta, jossa sen kehittelyyn käytetty GPS-aineisto on kerätty. On huomattava, että taulukoiden algoritmien toimivuudesta ja onnistumistarkkuudesta kertovat kohdat eivät ole yhteismitallisia, sillä onnistumista on eri tutkimuksissa mitattu hieman eri tavoin. Lopuksi arvioidaan kootusti algoritmien ja niiden taustalla olevien ideoiden hyödyllisyyttä oman algoritmin kehittämisessä.

1. Chungin ja Shalabyn (2005) algoritmi

Taulukko 6. Chungin ja Shalabyn (2005) kulkutavanpäättelyalgoritmin taustatiedot

Lähde	Chung ja Shalaby (2005)
Liikenne-ympäristö	Toronto (Kanadassa)
Lähtötiedot	GPS-aineistosta keskinopeus ja maksiminopeus, lisäksi paikkatietoaineistoja (tieverkko ja joukkoliikenteen reitit tutkimusalueella)
Aineiston validointi	GPS-aineisto koottiin valitsemalla oikean liikkumistutkimuksen tuloksista 60 matkaa, jotka toistettiin GPS-laitteella tallentaen ja tarkkaa matkapäiväkirjaa pitäen (lähtöpaikka, määräpaikka, reitti jne.) Matkojen todellinen kulutapa oli siis tiedossa.
Kulkutavat	Kulkutapa (oikein tunnistetut / ko. kulkutavalla tehdyt matkat) Kävely (23/23) Pyöräily (3/3) Auto (21/24) Bussi (8/10) Yhteensä (55/60) eli 92 % tunnistettiin oikein.
Muuta	Matkat sovitettiin aluksi tieverkon linkeille, ja tietoa bussireiteistä ja -pysäkeistä käytettiin hyväksi bussimatkojen tunnistamisessa.
Tulevaisuuden kehityssuunnat	Testaaminen laajemmalla aineistolla paljastaisi käytettyjen oletusten ja sääntöjen järkevyyden.

Algoritmin lähtöoletuksena on, että matkoilla on vain yksi tarkoitus ja matkalla käytetyt kulkutavat ovat joko pelkkä kävely, kävely-pyörä-kävely, kävely-auto-kävely tai kävely-bussi-kävely. Lisäksi oletetaan, että matka ei ala eikä lopu korkeiden rakennusten varjoon, jossa signaalin kuuluvuus on huono, eikä GPS-laitteen hidasta käynnistymistä (kylmä tai lämmin käynnistys) huomioida. Jotta nopeuden epätarkkuudet eivät vaikuttaisi liikaa, nopeusdataa tasoitetaan laskemalla jokaiselle pisteelle uusi nopeus keskiarvona kyseisen pisteen, edellisen pisteen ja seuraavan pisteen alkuperäisistä nopeuksista.

Kulikutavan vaihtoon tarvitaan siis kävelyä. Sen perusteella matkat jaetaan osiin, joiden välissä on kulkutavanvaihtopiste, jossa siis kulkutapa vaihtuu kävelystä joksikin muuksi tai päinvastoin. Kävely tunnistetaan sen matalasta nopeudesta.

Kulkutavanvaihtopisteet luokitellaan kolmeen luokkaan: kävelyn alku, kävelyn loppu ja aukon loppu (aukko tarkoittaa aikaa jolloin GPS-signaalia ei saada eli reittipisteitä ei tallennu). Pisteitä etsitään hakemalla kohtia, joissa nopeus on ensin alle 10 km/h ja sitten yli 10 km/h tai päinvastoin sekä kohtia, joissa signaali häviää.

Kun kulkutavanvaihtopisteet on löydetty, päätellään kulkutavat yksi kerrallaan. Jos kulkutavanvaihtopisteet ovat alle 40 metrin päässä bussipysäkestä, oletetaan että kulkutapa on bussi. Jäljelle jääneistä matkoista tarkastellaan maksiminopeutta: jos se on yli 32 km/h, niin kyseessä on auto. Näin jäljelle jääneistä matkoista ne, joiden maksiminopeus on yli 10 km/h, ovat pyörämatkoja ja loput ovat kävelymatkoja.

2. Bohten ja Maatin (2008) algoritmi

Taulukko 7. Bohten ja Maatin (2008) kulkutavanpäättelyalgoritmin taustatiedot

Lähde	Bohte ja Maat (2008)
Liikenneympäristö	Kolme erikokoista kaupunkia Hollannissa (n. 19 000, 61 000 ja 137 000 asukasta)
Lähtötiedot	GPS-aineistosta keskinopeus ja maksiminopeus, lisäksi käytetään paikkatietoja (rautateista).
Aineiston validointi	Keskityttiin erityisesti aineiston validointiin, joka tässä suoritettiin niin, että osallistujat saivat jälkikäteen nähdä matkansa internet-sovelluksessa ja heidän piti korjata väärät ja puuttuvat tiedot. Lisäksi tuloksia verrattiin Hollannin henkilöliikennetutkimuksen tuloksiin.
Kulkutavat	Kulkutapa (miten suuri osuus tunnistettiin, ja tutkittujen matkojen kokonaismäärä): Kävely (68 %, N=5481) Pyörä (72 %, N=8653) Auto (75 %, N=18017) Bussi/raitiovaunu/metro (0 %, N=328) Juna (34 %, N=747) Muu (7 %, N=460) Keskimäärin 70 % tunnistettiin oikein.
Mitä kulkutapoja vaikea erottaa toisistaan, miten algoritmi muuten toimi	Autoa ja juna on vaikea erottaa nopeuden perusteella, joten käytetään tietoa rautateiden ja rautatieasemien sijainnista; juoksijan kulkutavaksi tunnistetaan todennäköisesti pyöräily; mopoja ja muita vähemmän yleisiä kulkutapoja varten ei ole sääntöjä.
Tulevaisuuden kehityssuunnat	Nopeuden muutoksia analysoimalla pystytään mahdollisesti erottamaan useampia kulkutapoja.

Keskinopeus ja maksiminopeus saadaan GPS-aineistosta, mutta niiden perusteella ei vielä pysty varmasti päättelemään kulkutapaa, joten apuna tarvitaan myös paikkatietoja. Tässä algoritmissa käytetään lähinnä päätelystäntöjä.

Aluksi etsitään kävelymatkat: jos keskinopeus on alle 10 km/h ja maksiminopeus alle 14 km/h, asetetaan kulkutavaksi kävely. Sitten poimitaan jäljelle jääneistä pyöräilymatkat: jos keskinopeus on alle 25 km/h ja maksiminopeus alle 45 km/h, kyseessä on pyöräily. Näin jäljelle jääneiden matkojen keskinopeus tarkastetaan: jos se on alle 200 km/h, kyseessä on automatka.

Lopuksi etsitään erikseen junamatkoja paikkatiedon eli junaratojen sijainnin avulla. Jos piste on alle 50 metrin etäisyydellä rautatiestä, se määritellään ”ratapisteksi”. Jos jonkin matkan maksiminopeus on yli 20 km/h, matkassa on yli 20 ratapistettä ja ratapisteitä on yli puolet matkan kaikista pisteistä, asetetaan kulkutavaksi juna.

Lisäksi peräkkäiset junamatkat yhdistetään yhdeksi junamatkaksi. Koska GPS-signaali kuuluu usein huonosti junissa, etsitään myös mahdollisesti kokonaan puuttuvia junamatkoja. Jos siis jokin matka päättyy lähelle juna-asemaa ja seuraava alkaa sellaisen läheltä, ja välissä on kestoja on yli 3 minuuttia ja pituutta yli 5 km, oletetaan että väli on kuljettu junalla ja luodaan uusi junamatka.

3. Schuesslerin ja Axhausenin (2008) algoritmi

Taulukko 8. Schuesslerin ja Axhausenin kulkutavanpäättelyalgoritmin taustatiedot

Lähde	Schuessler ja Axhausen (2008)
Liikenneympäristö	Kolme kaupunkia Sveitsissä (n. 100 000, 185 000 ja 360 000 asukasta)
Lähtötiedot	GPS-aineistosta nopeuksien mediaani, nopeuksien 95. persentiili ja kiihtyvyyksien 95. persentiili. 95. Paikkatietoja ei käytetä.
Aineiston validointi	Tietoa todellisuudessa käytetyistä kulkutavoista ei ollut käytettävissä. Koska tietoa osallistujista ei ollut, mutta aineistoa oli paljon, validointi tehtiin vertaamalla tuloksia kyseisten kaupunkien otokseen vuoden 2005 liikkumistutkimuksessa (Swiss Microcensus on Travel Behaviour 2005)
Kulkutavat	Kulkutapa (osuus Swiss Microcensuksessa, osuus algoritmilla) Kävely (n. 55 %, n. 45 %) Pyörä (n. 5 %, n. 8 %) Auto (n. 20 %, n. 30 %) Kaupunkien julkinen liikenne (bussi ja raitiovaunu) (n. 15 %, n. 10 %) Juna (n. 4 %, n. 1 %) (Kokonaisnostumista ei voida arvioida, koska vertailutietoa oikeista kulkutavoista ei ollut.)
Muuta	Seurailee ainakin kulkutavanvaihtopisteiden käytön osalta metodia, jonka ovat tehneet Chung ja Shalaby (2005) ja Tsui ja Shalaby (2006).
Mitä kulkutapoja vaikea erottaa toisistaan, miten algoritmi muuten toimii	Junamatkoja saadaan esiin liian vähän.
Tulevaisuuden kehityssuunnat	Tulossa muun muassa kylmäkäynnistysongelman korjaus, puuttuvan datan interpolointi ja koko prosessin validointi aineistolla, jonka kulkutavat tunnetaan.

Aineiston esikäsittelyyn, suodattamiseen ja tasoittamiseen kiinnitetään paljon huomiota ennen kuin aineistoa ruvetaan jakamaan osiin ja soveltamaan siihen kulkutavan-tunnistusalgoritmia. GPS-pisteiden paikat tasoitetaan useamman niitä edeltävän ja seuraavan pisteen paikkojen perusteella, ja nopeudet ja kiihtyvyydet lasketaan tasoitettujen pisteiden perusteella.

Algoritmin lähtöoletuksena on, samoin kuin Chungin ja Shalabyn algoritmissa, että kulkutavan vaihtoon tarvitaan kävelyä. Kulkutavanvaihtopisteet etsitään samantyyllisesti kuin heidän algoritmissaan.

Sen jälkeen kun kulkutavanvaihtopisteet on löydetty, kulkutapojen päättelyyn käytetään sumeaa logiikkaa. Sumeina muuttujina ovat nopeuksien mediaani, nopeuksien 95.

persentiili ja kiihtyvyyksien 95. persentiili. 95. persentiiliä käytetään maksimiarvon sijaan, koska silloin poikkeavat havainnot eivät aiheuta niin paljon ongelmia.

Algoritmin onnistumistodennäköisyyttä ei voida suoraan laskea, koska tietoa todellisista kulkutavoista ei ole, mutta vertailu Sveitsin henkilöliikennetutkimuksen tuloksiin osoittaa, että algoritmin tuottamat tulokset ovat jossain määrin uskottavia. Tarkoituksena on tutkia algoritmin toimivuutta myös aineistolla, jonka kulkutavat tunnetaan.

4. Stopherin ym. (2008) algoritmi

Taulukko 9. Stopherin ym. (2008) kulkutavanpäättelyalgoritmin taustatiedot

Lähde	Stopher ym. (2008)
Liikenneympäristö	Australia, lähinnä suuria kaupunkeja ja niiden ympäristöjä (Adelaide, Canberra)
Lähtötiedot	GPS-aineistosta keskinopeus ja maksiminopeus, lisäksi käytetään taustatietoja vastaajista ja paikkatietoja.
Kulkutavat	Kävely Pyörä Auto Bussi Juna (Muut joukkoliikennevälineet) Keskimäärin 95 % kulkutavoista tunnistettiin oikein.
Muuta	Kylmäkäynnistysongelmaan on kehitetty ratkaisu, samoin korkeiden rakennusten välissä tapahtuvaan signaalin katoamiseen.
Mitä kulkutapoja vaikea erottaa toisistaan, miten algoritmi muuten toimii	Algoritmi toimii huonoiten silloin, kun GPS-aineisto on reikäistä, eikä kulkutapaa siksi voida tunnistaa varmasti. Etenkin bussimatkoja on vaikea erottaa automatkan, joilla signaali kuuluu huonosti tai jotka tehdään ruuhkaisissa oloissa.

Kulkutavanpäättelyssä käytetään hyväksi taustatietoja vastaajista ja paikkatieto-aineistoja. Aluksi päätellään todennäköisyysmatriisin ja painokerrointen avulla, onko kyseessä kävely-, pyöräily- vai moottoriajoneuvomatka. Jos taloudessa ei ole yhtään polkupyörää, pyöräilyn painoksi tulee nolla. Muiden painojen päättelyyn käytetään useita muuttujia: keskinopeus (lasketaan sekä keskinopeus=etäisyys/aika että GPS-laitteen tallentamien nopeuksien keskiarvo), maksiminopeus, useimmin käytetty nopeus (moodi), ja matkan pituus. Eri nopeusalueille annetaan eri painot, esimerkiksi kävely saa suuren painon jos keskinopeus on alle 6 km/h ja maksiminopeus alle 10 km/h.

Seuraavaksi erotellaan moottoroidulla ajoneuvolla tehdyt matkat toisistaan paikkatietoa apuna käyttäen. Ensin poimitaan joukkoliikennematkat: yleensä niitä ennen, niiden jälkeen tai sekä että on kävelymatka. Matkan alun ja lopun pitäisi osua joukkoliikennelinjalle, ja lisäksi voidaan yleensä havaita pysähdyksiä, jotka eivät osu risteyskiin. Joskus taas signaalin vastaanotto liikennevälineessä huono, joten jos kävelymatkan jälkeen jäljet häviävät ja jatkuvat jonkin matkan päässä, lasketaan tälle välille keskinopeus. Jos se on 10–40 km, kyseessä on luultavasti bussimatka. Jos taas nopeus on suurempi ja matkan alku tai loppu on rautatieaseman kohdalla, kyseessä on luultavasti junamatka. Lisäksi on O-bahn ja Adelaiden raitiovaunu, jotka voidaan helposti tunnistaa reiteistään, jos GPS-laite on tallentanut pisteitä matkan aikana.

Nyt pitäisi olla jäljellä vain polkupyörä- ja automatkoja, jotka erotetaan toisistaan lähinnä nopeuden perusteella. Jos keskinopeus on alle 40 km/h, kyseessä on yleensä polkupyörä, paitsi tunnetusti ruuhkaisilla alueilla. Lisäksi automatkat kulkevat yleensä katuverkkoa pitkin, kun taas pyörä voi poiketa siltä. Lisäksi tarkistetaan vielä, kulkeeko koko matka bussireittiä pitkin, ja jos näin on, niin tehdään vielä lisätutkimuksia muun muassa pysähdyksistä.

Arvio esitellyistä algoritmeista

Kaikki edellä esiteltyt algoritmit ovat vielä joiltain osin kehitysvaiheessa, ja tuovat hyvin esiin niitä ongelmia, joita datan jakamiseen matkoiksi tai matkan osiksi ja kulkutavan tunnistamiseen liittyy. Chungin ja Shalabyn algoritmin aineisto on hyvin pieni, vain 60 matkaa, joten hyvää tunnistamisprosenttia (92 %) ei voida pitää vielä kovinkaan luotettavana. Lähtöoletus, että kulkutavanvaihdokseen liittyy aina kävelyä, on hyvä alkuarvaus, mutta voi aiheuttaa ongelmia, kun matkan alku puuttuu.

Sekä Chungin ja Shalabyn että Bohten ja Maatin algoritmit perustuvat melko yksinkertaiseen päättelyyn keskinopeudesta, maksiminopeudesta ja joukkoliikenteen pysäkeistä. Bohten ja Maatin algoritmilla saavutetaan jo suhteellisen hyviä tuloksia (yhteensä 70 % oikein yli 33 000 matkasta) kävelyn, pyörän ja auton osalta, joilla tehtyjä matkoja aineistossa on paljon. Sen sijaan bussin ja junan tunnistaminen ei vielä toimi kunnolla.

Schuesslerin ja Axhausenin algoritmissa käytetään kehittyneempää päättelytapaa, sumeaa logiikkaa. Algoritmi näyttää toimivan varsin hyvin ja sumeaa logiikkaa käyttämällä saadaan intuitiivisesti mukaan useampia muuttujia. Etenkin junamatkojen tunnistamisessa on kuitenkin vielä tässä vaiheessa ongelmia, eikä ilman parempaa validointia pystytä sanomaan, sekoittuvatko jotkin kulkutavat toisiinsa. Muista esitellyistä algoritmeista poiketen tässä ei käytetä hyväksi paikkatietoa joukkoliikenteen pysäkeistä ja reiteistä, mikä on sikäli hyvä, että silloin ollaan riippumattomia saatavilla olevasta paikkatiedosta ja sen laadusta, mutta tulokset ovat jonkin verran huonompia (Schuessler ja Axhausen 2008).

Stopherin ym. algoritmissa taas käytetään todennäköisyysmatriiseja, painokertoimia ja monimutkaisia päättelysääntöjä. Tulokset ovat hyviä (95 %:ssa matkoista kulkutapa oikein) ja parametrit ovat ehtineet hioutua, mutta algoritmi kaikkine vaiheineen on jo melko monimutkainen.

Yhteenvetona oman algoritmin kehittämistä ajatellen voidaan sanoa, että melko yksinkertaisillakin päättelysäännöillä päästään jo suhteellisen hyviin tuloksiin ainakin kävelyn, pyöräilyn ja automatkojen osalta. Joukkoliikenteen tunnistaminen on haastavampaa, ja toisaalta joukkoliikenteen muodot vaihtelevat maittain ja kaupungeittain, joten yleismaailmallista tapaa eri joukkoliikennemuotojen erottamiseen tuskin on.

4 TUTKIMUKSEN TOTEUTUS

4.1 Kenttätutkimus

4.1.1 Laitteiden valinta ja hankkiminen

Laitteiden ominaisuuksista

GPS-laitteiden valinnassa oli otettava huomioon useita sekä kirjallisuusselvityksessä että hankintaprosessin aikana esiin tulleita asioita. Kirjallisuusselvityksen perusteella tultiin siihen tulokseen, että aluksi kannattaa määritellä, mitä ominaisuuksia laitteilta toivotaan, ja sitten etsiä mahdollisimman hyvin kriteerit täyttävä laite.

Kirjallisuustutkimuksen perusteella toivottavia ominaisuuksia laitteelle olivat:

- helppokäyttöinen ja helppo kantaa mukana -> vain yksi laite, joka sekä paikantaa että tallentaa, mahdollisimman pienikokoinen ja kevyt
- riittävä akun kesto (vähintään noin 16 h jolloin yöllä lataamisen pitäisi riittää)
- nopea käynnistyminen
- akku helppo ladata
- ei mene helposti itsestään pois päältä
- riittävästi tallennuskapasiteettia
- sopiva tallennusväli (korkeintaan 10 s) tai mahdollisuus säätää tallennusväliä
- tietojen tallennus standardinmukaisessa (NMEA) muodossa tai muuten selkeässä muodossa
- aineiston siirtäminen tietokoneelle mahdollisimman helppoa (esimerkiksi USB-portin tai Bluetooth-yhteyden kautta)

Erilaisia paikannus- ja navigointisovelluksia on kuluttajille myynnissä lukemattomia. Käytännössä keskityttiin etsimään mahdollisimman yksinkertaista ja helppokäyttöistä laitetta, joka täyttäisi vaaditun toiminnallisuuden, eli osaisi vastaanottaa GPS-signaalia, paikantaa itsensä ja tallentaa kuljetun reitin. Tallennusominaisuus, etenkin niin että laitteen muistiin mahtuu useamman päivän matkat, ei ole mitenkään itsestään selvä. Useat halvemmat laitteet ainoastaan vastaanottavat GPS-signaalin ja välittävät sen eteenpäin esimerkiksi navigointisovellukselle.

Oletuksena oli, että tarkoitukseen sopivia laitteita voisi saada halvimmillaan noin 50 eurolla. Budjetti oli tässäkin tutkimuksessa rajallinen, joten laitteiden etsinnässä kiinnitettiin erityistä huomiota hintaan. Toisaalta kalliimmissa laitteissa oli usein ominaisuuksia, jotka katsottiin tutkimuksen kannalta tarpeettomiksi (esimerkiksi retkeilykäyttöön tarkoitetuissa näyttö ja karttoja tallennettuna, ja GPS-sirulla varustetuissa älypuhelimissa lähes kaikki muut toiminnot). Rajaavaksi tekijäksi osoittautui hinnan ohella saatavuus. Laitteiden tilaamista ulkomailta ei tässä pidetty realistisena vaihtoehtona, vaan rajoituttiin kaupallisiin laitteisiin, jotka olivat välittömästi ostettavissa Suomesta tai suomalaisista verkkokaupoista. Suuremmassa tutkimuksessa suositeltavampi tapa olisi ehkä hankkia jo aiemmin liikkumistutkimuksiin soveltuviksi ja varmasti toimiviksi havaittuja laitteita, tai tehdä yhteistyötä laitevalmistajan kanssa.

Laitteiden hankinta käytännössä

Aluksi etsittiin mitä tahansa tarkoitukseen sopivia GPS-laitteita. Näistä valittiin tarkempaan vertailuun ne, jotka olivat hintaluokaltaan ja ominaisuuksiltaan sopivimmat, ja pystyivät tallentamaan tuhansia reittipisteitä. Vertailu on liitteessä 1. Laitteista pyrittiin selvittämään muun muassa akun kesto, tallennettavien reittipisteiden määrä, kuinka usein laite tallentaa pisteitä ja voiko tallennusväliä säätää, missä muodossa

laiteesta saa tallennetun tiedon ulos, saako tiedon suoraan näkymään karttapohjalla ja kuinka helposti laitteen voi ostaa.

Jotta lopullista valintaa ei olisi tarvinnut tehdä aivan sokkona, ostettiin kokeiltavaksi alustavasti hyvältä vaikuttava laite (i-Bluen valmistama iBT747 GPS Data Logger), jonka toimivuutta kokeiltiin noin kuukauden ajan. Kokeilulla selvitettiin muun muassa akun todellista kestoa, satelliittien löytymiseen kuluvaa aikaa ja signaalin kuuluvuutta riskioloissa, kuten joukkoliikennevälineissä. Samalla oli mahdollista kokeilla ja hioa myös vapaaehtoisten täytettäväksi suunniteltua matkapäiväkirjaa.

Laitetta kannettiin kokeilujen aikana takin taskussa ja repun sivutaskussa, joissa molemmissa se toimi hyvin. Tosin jos laitteen laittoi samaan taskuun matkakortin kanssa niin, että matkakortti peitti antennin, laite ei saanut yhteyttä satelliitteihin. Laite ei siis välttämättä toimi missä tahansa oloissa. Parhaat olosuhteet lienevät melko tyhjässä takin taskussa tai repun tai laukun taskussa, jossa laite on mahdollisimman lähellä laukun reunoja ja antennin ja satelliittien välissä on lähinnä kangasta. Missään vaiheessa laite ei mennyt pois päältä laukkuun tai taskuun laitteesta, mutta tiukkaan taskuun työntäminen voisi periaatteessa liikuttaa laitteen sivussa olevaa kytkintä niin, että laite menisi pois päältä.

Akun kesto todettiin riittäväksi; laitteen lataaminen yön yli mahdollisti sen pitämisen päällä koko seuraavan päivän. Hyvissä oloissa (kävelynopeus, riittävästi taivasta näkyvissä) laite löysi satelliitit ja aloitti paikantamisen noin puolessa minuutissa, kuten luvattu. Laite säilytti yhteyden satelliitteihin auto- ja bussimatkoilla hyvin ja junamatkoillakin suurimmalla osalla matkoista (signaali hävisi useamman kerran Pasilan asemalla, luultavasti aseman katosten takia).

iBT747 käyttää MTK-sirua, joten vertailun vuoksi kokeiltiin myös toisenlaisella sirulla (SIRF Star III) varustettua Globalsatin valmistamaa BT-335 Bluetooth Data Loggeria. BT-335 ei lyhyessä testissä osoittautunut signaalin löytymisen kannalta erityisesti paremmaksi tai huonommaksi kuin iBT747. Koska iBT747 oli hieman halvempi ja miellyttävämpi käyttää, ja sen toimivuutta oli testattu pitempään, päätettiin tutkimuskäyttöön ottaa iBT747. Lisäksi BT-335:n hankkimisessa edes testikäyttöön ilmeni yllättäviä saatavuusongelmia, eikä haluttu ottaa riskiä, että valitaan laite, jota ei olekaan saatavilla.

Vaikka saatavuusriskiä yritettiin välttää, siinä ei täysin onnistuttu. Siinä vaiheessa kun iBT747-laitteita piti ostaa erä tutkimusta varten, laite olikin korvattu uudemmalla versiolla, eikä edellistä, testattua versiota ollut enää saatavilla. Myyjä kuitenkin vakuutti, että uusi versio oli yhtä lisäominaisuutta lukuun ottamatta sama kuin edellinen versio, joten päädyttiin tilaamaan 10 kpl uutta versiota (iBT747A+). Lisäominaisuus uudessa versiossa on AGPS-tuki, mikä tässä tapauksessa tarkoittaa sitä, että laitteeseen voidaan internetistä ladata satelliittien lentoratatiedot etukäteen, jolloin satelliittien löytämiseen kuluvan ajan luvataan pienevän puolesta minuutista 15 sekuntiin. Toimiessaan tämä lisäominaisuus on erityisen hyödyllinen, koska se pienentää satelliittien etsimisestä johtuvaa puuttuvien reitinosien ongelmaa. Lisäksi myyntipakkauksesta oli muutoksessa jätetty pois verkkolaturi, mikä sinänsä on järkevää jätteensynnyn ehkäisyä. Tässä tutkimuksessa ei kuitenkaan voitu olettaa, että kaikki osallistujat pystyisivät lataamaan GPS-laitteen tietokoneen USB-portista tai autossa, joten laitteiden lisäksi oli hankittava erikseen verkkolaturit.

Eräs valitun laitteen huono ominaisuus on, että sen akku ei kestä käyttöä yli kymmenen asteen pakkasessa. Tässä tutkimuksessa se ei ollut ongelma, koska tutkimus toteutettiin Etelä-Suomessa, ja osallistumispäivillä ei ollut merkitystä, jolloin voitiin pyytää osallistujia siirtämään tutkimus seuraavaan päivään, jos alun perin valittuna päivänä oli yli kymmenen astetta pakkasta. Toisaalta mikään vertailluista laitteista (niistä joissa

käyttölämpötila ilmoitettiin) ei kestä yli 20 asteen pakkasta, mikä koko Suomen laajuisessa tutkimuksessa saattaisi muodostua ongelmaksi.

4.1.2 Kenttätutkimuksen suunnittelu

Yleistä tutkimuksen suunnittelusta

Kenttätutkimuksessa vapaaehtoiset tutkimushenkilöt kantoivat GPS-laitetta mukanaan joitakin päiviä ja täyttivät samalla paperista matkapäiväkirjaa. Matkapäiväkirjasta selvitettiin sitten, mitä kulkuvälinettä osallistujat olivat käyttäneet missäkin kohden GPS-laitteen tallentamaa aineistoa.

Lisäksi valitussa GPS-laitteissa on nappula, jota painamalla laite tallentaa normaalia eroavan reittipisteen. Osallistujia pyydettiin laitteen mukana pitämisen lisäksi painamaan laitteessa olevaa nappulaa aina, kun he aloittivat matkan tai tulivat perille. Kohdat, joissa nappulaa oli painettu, näkyisivät kartalla, ja niiden koordinaatit voitaisiin selvittää, ja niistä voisi olla apua matkojen lähtö- ja määräpaikkojen selvittämisessä.

Kenttätutkimuksen suunnittelu koostui tutkimuslomakkeiden suunnittelusta ja testaamisesta, osallistujien rekrytoinnista, aikataulun suunnittelusta ja ohjeistuksen kokoamisesta, joita on kuvattu jäljempänä.

Lisäksi piti päättää, kuinka usein aineistoa tallennetaan. Mahdollisimman tiheällä tallennusvälillä on se hyvä puoli, että saadaan talteen esimerkiksi kiihdytykset ja jarrutukset. Käytännössä sekunti oli pienin tässä tapauksessa järkevä aikaväli tallennuksille, ja useat kaupalliset GPS-laitteet voidaan asettaa tallentamaan sekunnin välein¹³. Sekunnin välein tallennetulla aineistolla voidaan myös simuloida harvemmalla tallennusvälillä tallennettua aineistoa käyttämällä vain osaa havainnoista.

Huono puoli on, että pieni tallennusväli lisää kerätyn aineiston määrää, mutta tässä tutkimuksessa koehenkilöiden määrä oli niin pieni, että tallennuskapasiteetti ei muodostanut ongelmaa. Koska sopivaksi osallistumisajaksi määritettiin kaksi päivää osallistujaa kohti, ja todettiin, että yöksi sammuttamalla GPS-laitteen muistiin mahtuisi kahden päivän aineisto sekunnin välein tallennettuna, päätettiin kerätä aineistoa niin tiheästi kuin mahdollista eli sekunnin välein.

Lomakkeet

Tutkimuslomakkeita tehtiin kolme kappaletta: matkapäiväkirja, taustatietolomake ja loppukysely. Lomakkeet ovat liitteenä 3. Tavoitteena oli tehdä mahdollisimman lyhyet lomakkeet, joilla kuitenkin saataisiin kaikki tarvittavat tiedot. Lomakkeisiin tuli kuitenkin lopulta melko paljon kysymyksiä, sillä tutkimusaineistosta haluttiin sellainen, että sitä voitaisiin hyödyntää myös mahdollisissa jatkotutkimuksissa. Lisätietojen kirjaaminen aiheutti vastaajille hieman lisävaivaa, mutta oli kokonaisuutena vähemmän vaivalloista kuin kokonaan uuden tutkimuksen toteuttaminen. Kysymyksiä käsittelevissä kohdissa on mainittu ideat, joita niiden avulla voitaisiin testata.

Taustatietolomakkeen pohjana olivat yleiset kysymykset (ikä, sukupuoli, asuinpaikka, työssäkäynti jne.) ja liikkumistottumuksia eri kulkutavoilla kartoittavat kysymykset, joista voidaan tutkia, voidaanko niiden vastauksia hyödyntää automaattisessa kulkutavanpäätelyssä. Lisäksi kysyttiin tietoja paikoista, joissa vastaaja käy usein, jotta voitaisiin myöhemmin mahdollisesti tutkia, auttavatko tällaiset taustatiedot selvittämään

¹³ Markkinoilla on jo kuluttajillekin laitteita, joilla voi tallentaa viisi kertaa sekunnissa, mutta se on liikkumistutkimuksissa jo ehkä hieman liiankin lyhyt aika, koska aineiston määrä kasvaa viisinkertaiseksi, mutta liikkumistutkimuksissa hyödyllinen informaatio ei luultavasti kasva samassa suhteessa.

matkojen tarkoituksia. Muun muassa Stopher ym. (2008) ovat käyttäneet erilaisia vastaajista kerättyjä taustatietoja algoritmien lähtötietoina.

Matkapäiväkirjan ensisijainen tarkoitus oli saada matkojen kulkutavoista tietoa, jota voidaan verrata GPS-aineistoon. Olennaisinta oli siis saada talteen jokaisen eri kulkutavalla kuljetun osuuden alkamis- ja päättymisaika sekä tietenkin kulkutapa. Siksi matkapäiväkirjassa pyydettiin kirjaamaan jokainen eri kulkutavalla tehty osuus omalle rivilleen. Koska matkapäiväkirjasta näin tuli varsin monimutkainen, siihen laitettiin ohjeiden lisäksi sivun mittainen esimerkki siitä, miten päiväkirjaa toivottiin täytettävän.

Matkapäiväkirjassa kerättiin myös tiedot matkojen lähtö- ja saapumisaikoista, saapumisaikkojen tyypeistä ja matkojen tarkoituksista, jotta samaa aineistoa voidaan tarvittaessa käyttää myös matkan tarkoituksen automaattisen tulkitsemisen kehittämiseen. Matkan tarkoitus ja saapumipaikan tyyppi pyydettiin valitsemaan valmiista vaihtoehdoista¹⁴.

Lisäksi kerättiin vielä tiedot matkan aikana vallinneista sääoloista, jotta voitaisiin tutkia, vaikuttaako sää esimerkiksi matkanopeuksiin eri kulkutavoilla (erityisesti kevyt liikenne). Säätietojen kirjaaminen yhdistettiin matkapäiväkirjaan, jotta sää tiedot on helppo liittää oikeaan matkaan, matkojen välillä muuttuva sää tulee huomioitua, eikä lomakkeita tullut enää lisää. Osallistujien kirjaamien säätietojen lisäksi pääkaupunkiseudun säätietoja kerättiin Forecan internet-palvelusta¹⁵, josta sai tiedot neljän edellisen päivän säätilasta. Säätietojen hyödyntäminen oikeassa tutkimuksessa voi kuitenkin olla haastavaa, sillä sää on hyvin paikallinen ilmiö ja siten vallinneen säätilan löytäminen kullekin GPS-aineiston matkalle on vaikeaa, koska mittauspisteitä on melko harvassa.

Loppukyselyn pääasiallinen tarkoitus oli kartoittaa osallistujien kokemuksia tutkimuksesta. Sen avulla muun muassa tutkittiin, miten kuormittavaksi osallistujat kokivat tutkimukseen osallistumisen ja millaisia ongelmia tutkimuksen aikana esiintyi. Lisäksi pyrittiin vertaamaan tutkimusajan käyttäytymistä taustatiedoissa annettuun oletettuun käytökseen, jotta selviäisi, voidaanko keskimääräistä käytöstä käyttää hyväksi yksittäisten päivien käytöksen tulkitsemisessa.

Loppukyselyn ensimmäisessä osassa kysyttiin, millä kulkutavoilla suurin osa tutkimuspäivien matkoista tehtiin, ja oliko osallistujilla auto käytössään tutkimuspäivinä. Näiden ideana on, että vaikka laajemmassa tutkimuksessa ihmiset eivät täyttäisi yksityiskohtaista matkapäiväkirjaa, heidät olisi ehkä mahdollista saada tutkimusjakson jälkeen täyttämään lyhyt lomake, jossa he kuvailisivat keskimääräistä matkustuskäyttäytymistään tutkimusjakson aikana. Jos osallistuja tutkimusjakson jälkeen ilmoittaisi esimerkiksi, että hän käytti suurimmalla osalla matkoistaan autoa, voitaisiin GPS-aineistoa tulkittaessa epäselvissä tapauksissa olettaa, että auto on todennäköinen vaihtoehto. Kysymykset olivat siis tämän idean testaamista varten.

Lisäksi kysyttiin työaikoja tutkimuspäivinä (ja taustatiedoissa yleisintä kulkutapaa työmatkalla). Tämän tarkoituksena oli, että voitaisiin testata sellaista matkojen tarkoituksen päättämiseen liittyvää ideaa, että niiden kohdalla, joiden työmatkoja ei voida tunnistaa työpaikan sijainnin perusteella, työmatkat voisi ehkä tunnistaa säännöllisten työaikojen ja työmatkalla käytetyn kulkutavan perusteella.

¹⁴ Vaihtoehdot olivat samat kuin vuosina YTV:n vuosina 2007 ja 2008 toteuttaman *Pääkaupunkiseudun ja sen työssäkäyntialueen liikennetutkimuksen* henkilöhaastatteluosiossa.

¹⁵ <http://foreca.fi/Suomi/tasmasaa/Espoo>, <http://foreca.fi/Suomi/tasmasaa/Vantaa>, <http://foreca.fi/Suomi/tasmasaa/Helsinki>

Loppukyselyn toinen osa keskittyi osallistujien kokemuksiin. Osallistujilta kysyttiin ensinnäkin mielipidettä tutkimuksen kestosta ja siitä, kuinka pitkään vastaavaan tutkimukseen he voisivat osallistua. Sen jälkeen kysyttiin, kuinka pitkään tutkimukseen he voisivat kuvitella osallistuvansa, jos tarvitsisi ainoastaan pitää GPS-laitetta mukana eikä täyttää matkapäiväkirjaa. Lisäksi kysyttiin, mikä tutkimuksessa oli rasittavinta, ja millaisia vaikeuksia tutkimuksen aikana esiintyi. Kaikkiin kohtiin sai kirjoittaa myös lisäselityksiä, ja lopussa oli vielä tilaa vapaalle palautteelle.

Osallistujien rekrytointi, otoksen kattavuus ja aikataulu

Kenttätutkimuksen osallistujamääräksi sovittiin 50 henkeä, määrä joka olisi suhteellisen helppo saada rekrytoitua. Osallistumispäivien määrästä tultiin siihen tulokseen, että kahden päivän ajalta voitaisiin ehkä saada yhdeltä henkilöltä erilaisia matkoja, mutta useamman päivän aikana matkoissa luultavasti esiintyisi huomattavasti toistoa. Lisäksi paperisen matkapäiväkirjan täyttämisen oletettiin olevan niin vaivalloista, ettei sitä jaksaisi tehdä kovin montaa päivää. Päätettiin siis, että jokainen osallistuja kantaisi laitetta mukana kaksi päivää. Eri viikonpäivien välillä on vaihtelua siinä, miten liikenne sujuu eri välineillä, joten todettiin, että olisi hyvä saada aineistoa kaikilta viikonpäiviltä.

Osallistujat rekrytoitiin lähinnä opiskelijoista ja työkavereista, jotta rekrytointi olisi mahdollisimman helppoa. Opiskelijat myös käyttävät melko kattavasti erilaisia kulkutapoja. Tavoitteena ei siis ollut saada koko väestöä edustavaa otosta, vaan kerätä suhteellisen helposti riittävästi aineistoa eri kulkutavoista. Otosta myös täydennettiin jonkin verran jo osallistuneiden tuttavilla, jolloin saatiin lisää vaihtelua osallistujien ikin ja matkojen määräpaikkoihin. Jo alun perin otettiin huomioon myös se mahdollisuus, että kaikista kulkutavoista ei välttämättä saada riittävästi havaintoja, eli varauduttiin täydentämään otosta joidenkin kulkutapojen osalta.

Kenttätutkimukseen kuluva aikaa oli hankala arvioida etukäteen, mutta minimiajaksi osallistujaa kohti oletettiin viikko, johon siis sisältyisi laitteen toimitus, kaksi tutkimuspäivää ja laitteen palautus. Tavoitteeksi asetettiin saada kenttätutkimus suoritettua kuudessa viikossa.

Osallistujien ohjeistus

Kenttätutkimuksen osallistujille annettu kirjallinen ohjeistus on liitteessä 3 matkapäiväkirjan ensimmäisellä sivulla sekä liitteessä 4, jossa on ohjeistus tutkimuksen suorittamisesta ja GPS-laitteen käyttämisestä. Sama ohjeistus hieman suppeammin kerrottiin suurimmalle osalle osallistujista samalla kun GPS-laite ja lomakkeet annettiin heille. Samalla käytiin myös läpi laitteen toiminta.

Ohjeet pyrittiin laatimaan niin perusteellisesti, että tutkimuksesta oli tarvittaessa mahdollista suoriutua pelkästään kirjallisten ohjeiden perusteella. Samalla oli kuitenkin pidettävä mielessä, että osa ihmisistä lukee mielellään kaikki ohjeet tarkasti ennen tutkimuksiin osallistumista, kun taas toiset ihmiset eivät lue ohjeita yhtään enempää kuin on pakko. Ohjeiden laadinnassa oli tasapainoteltava näiden ääripäiden välillä, jotta ohjeita olisi riittävästi, mutta ei liika.

Ohjeistuksen tarkoituksena oli kertoa, miten laitetta oli suositeltavaa käyttää ja matkapäiväkirjaa täyttää, jotta tutkimukseen saataisiin mahdollisimman tarkoituksenmukaista ja täydellistä tietoa. Lisäksi ohjeessa oli tietoa laitteen rajoituksista, jotta osallistujat esimerkiksi tietäisivät, että laitteen toiminta sisätiloissa ei vaikuta tutkimuksen onnistumiseen. Osa ohjeista on GPS-laitteesta ja tutkimuksen luonteesta riippuvaisia, joten koko ohjetta ei käydä tässä läpi, vaan keskitytään muutamaa olennaiseen asiaan.

Jotta osallistujat eivät olisi unohtaneet laittaa GPS-laitetta päälle joidenkin matkojen ajaksi, heitä pyydettiin pitämään laitetta päällä koko päivä. Tarkemmin sanottuna, heitä pyydettiin laittamaan laite päälle aamulla lähtiessään kotoa, ja laittamaan se pois päältä ja latautumaan tullessaan illalla kotiin. Ohjeessa myös todettiin, että jos osallistuja ei halua jonkin matkan tallentuvan, hän voi laittaa laitteen pois päältä kyseisen matkan ajaksi, tai olla ottamatta laitetta mukaan. Vaikka erillistä lupaa joidenkin matkojen poisjättämiseen ei olisikaan annettu, olisi kuka tahansa kuitenkin osannut jättää laitteen kotiin, joten luvan antaminen toi näkyväksi vallitsevan käytännön.

Osallistujia ei erikseen ohjeistettu odottamaan liikkeelle lähtiessä, että GPS-satelliitit löytyvät, koska sen olisi koettu rajoittavan liikaa heidän liikkumistaan. Ei ole realistista odottaa, että laajemmassa tutkimuksessa ihmiset jaksaisivat odotella satelliittien löytymistä edes puolta minuuttia. Jos käytössä olisi samalla esimerkiksi elektroninen matkapäiväkirja, johon matkan alkaessa pitää kirjata matkan ominaisuuksia, satelliittien odottelu hoituisi kirjaamisen aikana esimerkiksi autossa. Sen sijaan esimerkiksi kävelymatkan tietoja saatetaan kirjata sisällä ennen lähtöä, jolloin satelliitteja ei ehditä löytää ennen liikkeellelähtöä.

Ohjeissa oli myös kerrottu, miten toimia jos GPS-laite näytti toimivan huonosti. Pahempia ongelmatilanteita ja kysymyksiä varten oli puhelinnumero, johon sai soittaa. Ohjeistusta myös laajennettiin ja muokattiin jonkin verran ensimmäisten osallistujien antaman palautteen perusteella.

4.1.3 Kenttätutkimuksen suorittaminen

Sen jälkeen kun vapaaehtoiset kenttätutkimukseen oli saatu rekrytoitua, tutkimus muodostui seuraavista osista:

- laite ja lomakkeet toimitettiin osallistujalle
- osallistuja täytti lomakkeet ja keräsi GPS-aineiston
- osallistuja toimitti laitteen ja lomakkeet takaisin (yleensä ne haettiin häneltä sopivasta paikasta)
- GPS-aineisto purettiin tietokoneelle ja tarkistettiin, GPS-laite tyhjennettiin seuraavaa käyttäjää varten
- lomakkeiden tiedot kirjattiin sähköiseen muotoon

Suurimmalle osalle osallistujista GPS-laite ja lomakkeet toimitettiin henkilökohtaisesti kotiin, työpaikalla tai johonkin muuhun sopivaan paikkaan. Samalla annettiin ohjeistusta lomakkeiden täyttämiseen ja laitteiden käyttämiseen. Muutama osallistuja ei saanut suullista ohjeistusta kiireen vuoksi, eli he pääsivät kokeilemaan, selvisikö tutkimuksesta pelkän kirjallisen ohjeistuksen varassa; he selvisivät yhtä hyvin kuin muutkin. Muutamalle osallistujalle laitteen ja lomakkeet toimitti henkilö, joka oli ensin itse vapaaehtoisena osallistunut tutkimukseen ja pystyi siis kertomaan, miten tutkimuksessa tuli toimia.

Tutkimuksen aikana osallistujat täyttivät taustatietolomakkeen ja pitivät GPS-laitetta mukanaan vähintään kaksi päivää täyttäen samalla matkapäiväkirjaa. Lisäksi osallistujien piti painaa GPS-laitteen päällä olevaa nappulaa aina kun he aloittivat matkan tai tulivat perille. Tutkimuspäivien päätteeksi osallistujat täyttivät loppukyselyyn kokemuksiaan tutkimuksesta.

Osallistujille ei ollut annettu tiettyjä päiviä, joina heidän pitäisi pitää laitetta mukana, vaan he saattoivat valita itselleen sopivat kaksi päivää. Näin osallistumiseen saatiin hieman lisää joustavuutta. Matkoihin saatiin myös lisää vaihtelua korostamalla, että olisi tutkimuksen kannalta hyödyllistä valita päiviä, joiden matkojen jo etukäteen tiedetään poikkeavan toisistaan.

Kun osallistujat olivat saaneet kaksi tutkimuspäivää täyteen, heidän kanssaan sovittiin palauttamiselle aika ja paikka. Kun GPS-laite oli saatu takaisin, siihen tallennettu aineisto purettiin tietokoneelle ja tarkistettiin, että se näytti järkevältä. Sama aineisto tallennettiin useammassa muodossa, jotta sitä voitaisiin tarkastella ja muokata eri ohjelmilla. Sitten GPS-laitteen muisti tyhjennettiin seuraavaa käyttäjään varten, ja laitteeseen ladattiin AGPS-päivitys (tiedot satelliittien tulevista liikkeistä noin kuudeksi päiväksi). Samoin tarkastettiin, että lomakkeet oli täytetty, ja lomakkeiden tiedot kirjattiin sähköiseen muotoon.

4.2 GPS-aineiston käsittely ja analysointi

Aineiston käsittely

Tietoaaineistojen käsittelyllä tarkoitetaan tässä GPS-aineiston muokkaamista yhdessä tiedostossa olevasta jatkuvasta pisteiden jonosta erillisiksi tiedostoiksi, joissa jokaisessa on yhdellä kulkutavalla tehty matkan osuus. Muokkaus tehtiin poimimalla GPS-aineistosta paperisessa matkapäiväkirjassa mainitut matkat seuraavassa kuvatulla tavalla.

Kun GPS-aineistot oli purettu tietokoneelle, ne tallennettiin tekstimuodossa csv-tiedostoksi, jota on helppo jatkokäsitellä erilaisilla ohjelmilla. Matkapäiväkirjoista puolestaan poimittiin jokaisen eri kulkutavalla tehdyn matkan osan päivämäärä, lähtöaika, saapumisaika ja kulkutapa. Samalla jokaiselle henkilölle annettiin yksilöllinen tunniste, samoin jokaiselle matkalle ja matkan osalle. Prosessissa jouduttiin välillä tekemään tulkintoja siitä, miten yksi matka määritellään. Matkapäiväkirjoissa oli myös hieman puutteita, joten välillä jouduttiin tarkistamaan kellonaikoja GPS-aineistosta tai jättämään joitain kohtia tyhjiksi. Aivan kaikkia matkapäiväkirjoissa olleita matkoja ei löytynyt GPS-aineistosta, koska muutamassa tapauksessa laitteen muisti oli täyttynyt, kun laite oli jäänyt yöksi päälle.

Seuraavaksi GPS-aineistoa käsiteltiin automaattisesti käyttäen GPSBabel-ohjelmaa¹⁶. Ensin UTC-ajassa olevia GPS-pisteitä siirrettiin eteenpäin 2 tuntia (talviaika) tai 3 tuntia (kesäaika), jotta ne saatiin vastaamaan matkapäiväkirjoissa käytettyä paikallista aikaa. Päivämäärien ja kellonaikojen perusteella poimittiin GPS-aineistosta edelleen GPSBabel-ohjelmalla jokaiseen matkan osaan kuuluvat pisteet omaan tiedostoonsa.

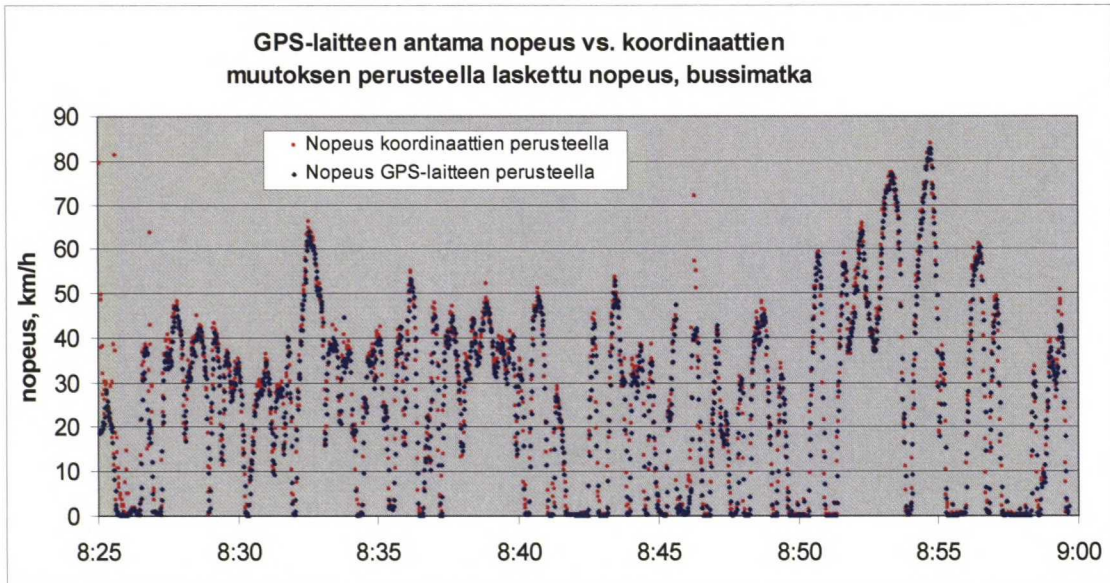
Tässä prosessissa aiheutui jonkin verran virhettä siitä, että käytettiin matkapäiväkirjassa ilmoitettuja kellonaikoja minuutin tarkkuudella, mutta osallistujan kellon ja GPS-laitteen kellonajan välillä voi olla eroa. Eroa on jo siksi, että ihmisten kellot voivat helposti olla pari minuuttia edellä tai jäljessä, ja lisäksi osallistujat eivät välttämättä olleet aina muistaneet katsoa kelloa, vaan olivat arvioineet ajan jälkikäteen. Pitkillä matkoilla tämä virhe ei ole ongelma, mutta lyhyillä, etenkin alle 10 minuutin matkoilla, ongelma on huomattava, eikä sitä oikeastaan voi korjata muuten kuin käsin (tarkistamalla aineistosta). Ongelmia aiheutuu etenkin kävelymatkoilla, jotka ovat tyypillisesti lyhyitä, ja jotka alkavat usein heti liikennevälineestä poistumisen jälkeen, jolloin esimerkiksi oletetussa kolmen minuutin kävelymatkassa voikin olla oikeasti kaksi minuuttia bussimatkaa ja yksi minuutti kävelyä.

Analysointi

Lopuksi saatuja matkan osia analysoitiin hieman. Arvio nopeudelle jokaisessa pisteessä saatiin suoraan GPS-aineistosta. Pisteiden aika- ja nopeuserojen perusteella laskettiin arvio kiihtyvyydelle jokaisessa pisteessä. Nopeus ja kiihtyvyys olisi voitu määrittää

¹⁶ Lisätietoja GPSBabelista, <http://www.gpsbabel.org/>

myös koordinaattien muutoksen perusteella. Kun verrattiin eri tavoin laskettuja nopeuksia, havaittiin, että ne osuvat pääosin samalle käyrälle, vaikkakin pisteet ovat hieman eri kohdilla käyrää¹⁷. Nopeudet eräältä bussimatkalta on esitetty kuvassa 9. Liitteessä 5 on vertailtu nopeuksia ja kiihtyvyyksiä myös kävely- ja automatkoilla. GPS-laitteen laskemat nopeudet pysyvät paremmin yhtenäisellä käyrällä, kun taas koordinaateista lasketuissa nopeuksissa on poikkeavia pisteitä käyrän ulkopuolella. Käyrälle kuulumattomat nopeuspisteet voivat johtua yksinkertaisesti siitä, että ne on laskettu koordinaateista, joissa on ollut jonkin verran virhettä. Tätä tukee se, että epärealistisia pisteitä on esimerkiksi kuvassa 9 erityisesti alussa, jossa GPS-laite on juuri käynnistetty.



Kuva 9. Erot GPS-laitteen määrittämän nopeuden (siniset pisteet) ja koordinaattien muutoksesta lasketun nopeuden (punaiset pisteet) välillä eräällä bussimatkalla.

Tämän perusteella GPS-laitteesta suoraan saadut nopeudet ovat luultavasti luotettavampia, ja työssä on käytetty niitä. Jos haluttaisiin käyttää koordinaateista laskettua nopeutta ja kiihtyvyyttä (esimerkiksi jos GPS-laite ei määritä ja tallenna nopeutta), voitaisiin käyttää ensin koordinaattien tasoitustekniikkaa, jolloin satunnaiset virheet koordinaateissa pienenevät (kuten Schuessler ja Axhausen (2008) tekivät).

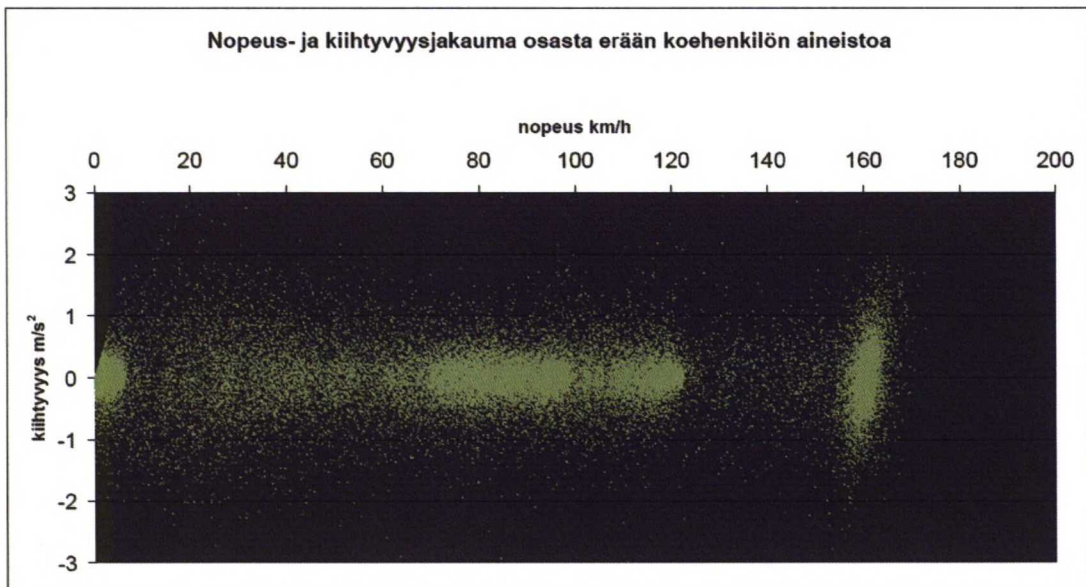
Koska kiihtyvyyttä esiintyi sekä negatiivisena että positiivisena, kiihtyvyyden keskiarvo on usein lähellä nollaa, joten laskettiin myös kiihtyvyyden itseisarvo jokaisessa pisteessä. Kaikille kolmelle tunnusluvulle (nopeus, kiihtyvyys, kiihtyvyyden itseisarvo joka pisteessä) laskettiin kultakin matkanosalta 7 tunnuslukua: keskiarvo, mediaani, maksimi, 90. persentiili, 95. persentiili, varianssi ja keskihajonta.

Eri kulkutapojen tunnuslukuja analysoitiin jonkin verran, jotta selviäisi, mitkä tunnusluvut erottelevat kulkutapoja toisistaan parhaiten. On huomattava, että sellainen tunnusluku, joka erottaa kaksi kulkutapaa toisistaan, ei välttämättä erota niitä kolmannelta. Kannattaa siis tutkia, mitkä kulkutavat sekoittuvat helposti toisiinsa, ja etsiä sellainen tunnusluku, joka erottaa juuri ne kaksi toisistaan (esimerkiksi auton laaja nopeusalue pakottaa etsimään muita keinoja erottaa se muista kulkutavoista).

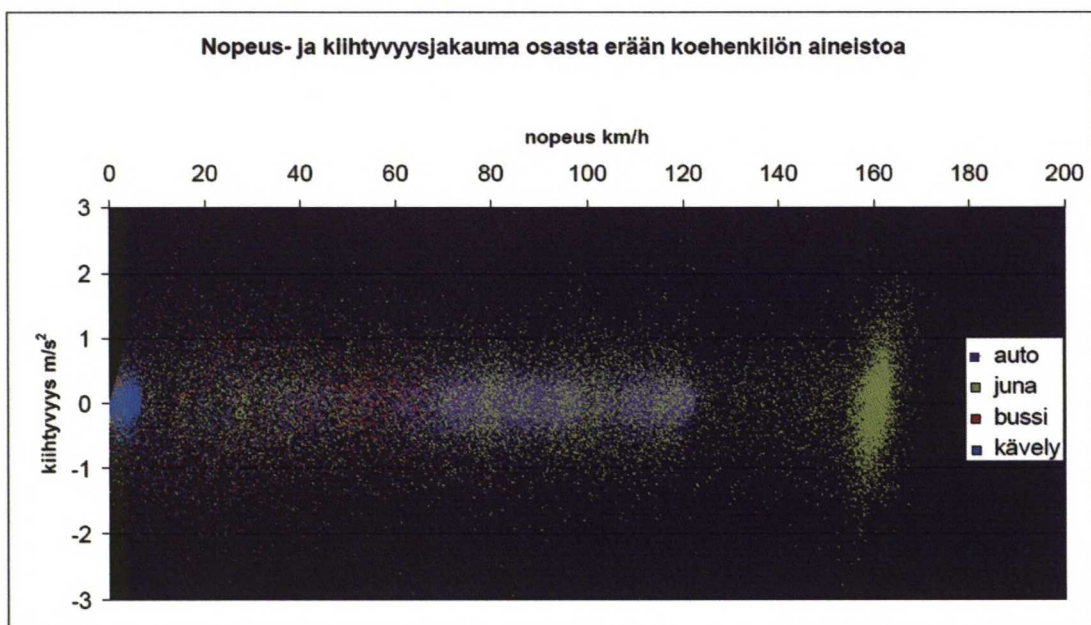
¹⁷ Pisteet ovat eri kohdilla luultavasti siksi, että GPS-laite määrittää nopeuden nimenomaan sillä hetkellä, kun piste tallennetaan. Jälkikäteen koordinaateista laskettu nopeus taas on tässä laskettu keskinopeutena edellisen sekunnin ajalta, ja se annetaan nopeudeksi pisteelle, joka on tuon sekunnin loppupäässä. Kiihdytyksessä siis GPS-laitteen antama nopeus on suurempi, ja hidastuksessa vastaavasti pienempi kuin laskettu nopeus.

Tunnuslukuja tutkittiin muun muassa taulukkolaskentaohjelmalla piirtämällä kuvaajia, tarkastelemalla laskettujen tunnuslukujen tilastollisia tunnuslukuja kulkutavoittain ja tutkimalla regressiomalleilla (multinomiaalinen logistinen regressiomalli), mitkä tunnusluvut selittivät eri kulkutapoja parhaiten.

Schuessler ja Axhausen (2008) käyttivät apuna kulkutapojen tunnistamisessa apuna seuraavan kuvan tyyppisiä nopeus- ja kiihtyvyyssjakaumakuvia (kuva 10). Heillä ei ollut tietoa siitä, mitä kulkutapaa mikin piste edusti, vaan he joutuivat päättämään, lähinnä nopeuden perusteella, mitä kulkutapoja eri pisteklusterit voisivat olla. Kuvassa 11 on sama jakauma kuin kuvassa 10, mutta siinä eri kulkutavat on eroteltu toisistaan. Kuva osoittaa hyvin, miksi pelkästään nopeuksia ja tarkastelemalla on vaikea erottaa kävelyä nopeampia kulkutapoja toisistaan luotettavasti pelkän nopeuden avulla.



Kuva 10. Esimerkki nopeus- ja kiihtyvyyssjakaumasta erään kenttätutkimukseen osallistuneen keräämässä aineistossa. Schuessler ja Axhausen käyttivät analyysissään samanlaisia kuvia. Tämä jakauma on tässä työssä kerätystä aineistosta.



Kuva 11. Sama nopeus- ja kiihtyvyyssjakauma kuin edellisessä kuvassa, mutta nyt on mukana tieto siitä, mitä kulkutapaa mikin piste edustaa. Nopeusalueella 80–120 km on junaa merkitsevien pisteiden alla piilossa vähintään yhtä paljon autoa merkitseviä pisteitä.

4.3 Algoritminkehitys

4.3.1 Algoritmien valinta

Algoritmien valinnan taustaa

Kuten kohdassa 2 (Käytetyt menetelmät ja työn rajaus) on todettu, työssä oli tavoitteena kehittää algoritmi eri kulkutapojen tunnistamiseen. Sen tueksi kehiteltiin alustavaa algoritmia myös GPS-aineiston jakamiseksi yhdellä kulkutavoilla kuljetuiksi matkan osiksi.

Muut, esimerkiksi tarkan reitin tai matkan tarkoituksen päättelymiseen tarkoitettut, algoritmit rajattiin kokonaan työn ulkopuolelle.

Lähtökohtana kulkutavantunnistus- ja aineistonjakamisalgoritmeille olivat kirjallisuusosiossa käsitellyt algoritmit. Niiden, omien ajatusten ja ohjausryhmän ideoiden pohjalta lähdettiin kehittämään ensin mahdollisimman yksinkertaista kulkutavantunnistusalgoritmia, johon sitten lisättiin muita lähtötietoja.

Kulkutavoista yritettiin tunnistaa kävelyn, pyöräilyn ja autoilun lisäksi erikseen bussi (sisältäen pitkänmatkan linja-autot), juna, raitiovaunu ja metro. Joukkoliikennevälineet voitaisiin niputtaa myös yhdistelmiksi, mutta koska kaikkien pysäkit ovat erikseen tiedossa, haluttiin katsoa, miten hyvin mikäkin niistä saadaan tunnistettua.

Kulkutavantunnistusalgoritmin lähtötiedot

Edellä todettiin, että kulkutavan tunnistamiseen voidaan käyttää ainakin GPS-aineistoa (tunnusluvut nopeuksista ja kiihtyvyyksistä), erityyppisiä paikkatietoaineistoja (joukkoliikenteen pysäkit ja asemat, joukkoliikenteen reitit ja tiet, kevyen liikenteen väylät) ja taustatietoja vastaajista.

Jos käytetään pelkkää GPS-aineistoa, voidaan käyttää hyväksi tietoa eri kulkutapojen nopeuksien ja kiihtyvyyksien jakaumasta ja vaihtelusta matkan aikana, kutsuttakoon tätä tietoa vaikkapa kulkutavan profiiliksi (esimerkiksi kuvassa 9 on erään bussimatkan nopeusprofiili). Kevyellä liikenteellä profiilit ovat melko universaaleja (ihmisen fysiikka asettaa rajat nopeudelle ja kiihtyvyydelle) ja vaihtelut johtuvat olosuhteista, kuten päällysteestä, märkisyydestä ja ehkä säästä. Joukkoliikennevälineiden profiilit taas vaihtelevat sen mukaan, millaiseksi kulkuneuvo ja väylä on milloinkin rakennettu (pysäkkiväli, oma väylä vai muun liikenteen seassa, kulkuneuvon rakenne). Tietoja eri välineiden profiileista saatiin tässä työssä kenttätutkimusaineistosta. Sen lisäksi käytettiin tunnettua tietoa eri kulkutapojen tyypillisistä ja korkeimmista nopeuksista.

GPS-aineistosta voidaan laskea myös tehtyjen matkojen osien kestoja ja pituuksia. Esimerkiksi henkilöliikennetutkimuksista saadaan tietoa eri kulkutavoilla tehtyjen matkojen keston ja pituuden jakaumasta, ja tätäkin tietoa voitaisiin ehkä käyttää kulkutapaa pääteltäessä. Vaikkeutena on kuitenkin tällä hetkellä se, että algoritmi analysoi matkan osia, jotka koostuvat yhdestä kulkutavasta. Valtakunnallisessa henkilöliikennetutkimuksessa taas on ilmoitettu kesto ja pituus vain koko matkalle, vaikka matkalta tiedetäänkin kaikki käytetyt kulkutavat (Henkilöliikennetutkimus 2004–2005). Pitäisi siis kerätä aineistoa, jossa on kysyttynä kaikkien matkan osien kesto ja pituus erikseen (mikä olisi hyvin työlästä), tai tehdä algoritmista sellainen, että vertailua tunnettuihin pituus- ja kestojakauksiin tehdään vasta, kun matkan osat on jo yhdistelty kokonaisiksi matkoiksi. Tällöin hankaluutena on, että osien yhdistely matkoiksi tapahtuu todennäköisesti jo tunnistettujen kulkutapojen perusteella (esimerkiksi ketju kävely-bussi-kävely yhdistetään yhdeksi matkaksi, jonka pääkulkutapa on bussi).

Pyörä- ja automatkoilla, jotka tyypillisesti koostuvat vain yhdestä kulkutavasta (tai kävelyosuus on niin lyhyt, ettei sillä ole juurikaan merkitystä), pituuksia ja kestoja voisi

ehkä käyttää jo nykyisillä aineistoilla. Tässä työssä näitä tietoja ei kuitenkaan käytetty, koska GPS-aineistosta lasketut matkan keston jakaumat eri kulkutavoilla poikkesivat niin paljon valtakunnallisen henkilöliikennetutkimuksen jakaumista¹⁸ (Henkilöliikennetutkimus 2004–2005). Poikkeama johtuu luultavasti siitä, että tässä tutkimuksessa oli niin vähän ihmisiä, että tulos ei ole tilastollisesti edustava, ja lisäksi siitä, että valtakunnallisessa henkilöliikennetutkimuksessa on koko Suomen asukkaiden matkat, mutta tässä tutkimuksessa mukana oli lähinnä pääkaupunkiseudun asukkaita. Jos kesto- ja pituusjakaumat havaittaisiin hyödyllisiksi, olisi luultavasti mahdollista kiertää ongelma poimimalla henkilöliikennetutkimuksen aineistosta ainoastaan pääkaupunkiseudun asukkaiden matkat tai käyttämällä YTV:n tekemien, vain pääkaupunkiseutua ja sen ympäristöä koskevien tutkimusten tuloksia. Matkojen pituuksien jakaumaa aineistossa ei edes tutkittu, koska matkan pituuden määrittäminen GPS-aineistosta on työläämpää kuin keston, ja ongelmat olisivat luultavasti samat kuin keston jakauman kanssa. Saatu lisähyöty kulkutavantunnistuksessa ei todennäköisesti olisi kovin suuri, vaikka tietoa voitaisiinkin käyttää.

Matkojen kestojen ja pituuksien käyttämisessä algoritmista on vaikeutena myös se, että esimerkiksi henkilöliikennetutkimuksissa matkojen kestojen ja pituuksien jakaumat ovat tutkimuksissa kerättävää tietoa. Vanhojen tulosten käyttäminen uusien johtamisessa voisi pahimmillaan johtaa siihen, että muutokset edellisiin vuosiin verrattuna jäisivät huomaamatta. Luultavasti on hyödyllisempää jättää kesto- ja pituusjakaumat vertailuaineistoksi, jonka perusteella voidaan päätellä, miten hyvin algoritmit toimivat, kuten esimerkiksi Schuessler ja Axhausen (2008) ovat tehneet. Tällöin algoritmilla analysoidut matkan osat voidaan ensin yhdistää matkoiksi, joita sitten vertaillaan.

Paikkatietoaineistojen käyttämisessä päättelyn apuna on hyvät ja huonot puolensa. Tiedot joukkoliikenteen pysäkeistä ja reiteistä voivat helpottaa kulkutapojen tunnistamista huomattavasti, etenkin jos samoilla väylillä voi liikkua niin polkupyöriä, autoja kuin joukkoliikennevälineitäkin suunnilleen samaa vauhtia. Toisaalta algoritmin rakentaminen sen varaan, että nämä tiedot ovat saatavilla, heikentää algoritmin siirrettävyyttä sellaisille alueille, joilta tietoa ei ole kerätty.

Koska pistemäisten koordinaattitietojen vertaaminen on helppoa, ja tiedot bussipysäkeistä ja juna-asemista koko Suomessa on ainakin jollain tarkkuudella kerätty DigiStop-järjestelmään, päätettiin käyttää hyväksi tietoa joukkoliikenteen pysäkkien paikoista¹⁹. Pääkaupunkiseudulla on lisäksi saatavilla raitiovaunupysäkkien ja metroasemien koordinaatit²⁰. GPS-aineiston vertaaminen joukkoliikenteen reitteihin on jo vaikeampaa, joten vaikka tiedot jossain muodossa ovat olemassa, niitä ei tässä työssä käytetä.

Osallistujiin liittyvien taustatietojen käyttämisessä kulkutavanpäättelyalgoritmissa on myös hyviä ja huonoja puolia. Esimerkiksi Stopher ym. (2008) käyttävät hyväksi tietoa siitä, että kotitaloudessa ei ole yhtään polkupyörää. Tieto siitä, että tiettyä matkaa ei polkupyörän omistamattomuuden takia kovin todennäköisesti ole tehty polkupyörällä, voi olla ratkaiseva, jos kulkutapaa muuten on vaikea erottaa vaikkapa autosta.

¹⁸ http://www.hlt.fi/tulokset/6_33_tapa.xls

¹⁹ Algoritmin toteutusta varten bussipysäkkien ja juna-asemien koordinaatit otettiin Digi-Stopista. Juna-asemien osalta tietoa täydennettiin YTV:n aineistolla, ja pitkille asemille etsittiin lisäksi käsin lisäpisteitä, koska yksi piste kuvaa suurta asemaa huonosti.

²⁰ Raitiovaunupysäkkien ja metroasemien koordinaatit saatiin joukkoliikennerekisteri JORE:sta YTV:ltä/HKL:ltä. Metroasemille etsittiin lisäksi käsin lisäkoordinaatteja, koska yksi piste kuvaa huonosti sitä, että metroasemat ovat pitkiä, ja monille metroasemille on tyypillisesti useita sisäänkäyntejä (GPS-signaali voi hävitä jo sisäänkäynnin kohdalla).

Esimerkiksi auton tai ajokortin omistamisesta ei kuitenkaan voida yhtä helposti tehdä johtopäätöksiä, sillä autoa todennäköisesti lainataan useammin kuin polkupyörää, ja usein autottomat saavat autokyydin. Kahden aikuisen kotitaloudessa auto voi lisäksi olla kumman tahansa käytössä.

Niin kauan kuin taustatietoja kerätään tutkimusten yhteydessä joka tapauksessa muita tarkoituksia varten, niiden käyttämisestä ei liene haittaa, jos niistä saadaan merkitsevää lisähyötyä. Tulevaisuudessa voi kuitenkin olla mahdollista, että esimerkiksi tarvittavat demografiset tiedot saadaan valmiista rekisteristä tai eri rekistereitä yhdistelemällä, jolloin kynnys aiheuttaa vastaajalle lisäkuormitusta kysymällä erikseen algoritmiin tarvittavia lisätietoja kasvaa. Toisaalta voi olla tarpeen kysyä joitain taustatietoja muita algoritmeja varten, esimerkiksi matkan tarkoituksen päättelyyn. Voi myös käydä niin, että tietosuojasysteemit olemassa olevaakaan tietoa ei saa yhdistellä. Taustatietojen keräämistä on siis joka tapauksessa mietittävä osana suurempaa kokonaisuutta.

Kulikutavantunnistusalgoritmin metodit

Kulikutavantunnistukseen päätettiin siis käyttää GPS-aineistosta saatavia nopeus- ja kiihtyvyydestietoja sekä tietoa joukkoliikenteen pysäkeistä. Joukkoliikenteen pysäkkien pistemäisyyttä korjattiin siten, että bussi- ja raitiovaunupysäkkien ympärille tehtiin laatikko, joka oli noin 50 metriä joka suuntaan pysäkkiä merkitsevästä pisteestä, ja juna- ja metroasemilla 100 metriä.

Nopeuden tunnusluvusta käyttöön valittiin keskinopeus, nopeuden 90. persentiili ja nopeuden keskihajonta. 90. persentiili valittiin, koska sitä käytettäessä väärit tai huonot havainnot aiheuttavat vähemmän ongelmia kuin jos käytettäisiin maksiminopeutta. Keskihajonta taas on havainnollisempi kuin varianssi, eikä lopputuloksen kannalta ole suurta väliä, kumpaa käytetään. Kuljetun matkan keskinopeus saadaan laskemalla sekunnin välein ilmoitettujen nopeuksien keskiarvo²¹.

Kiihtyvyyden tunnusluvusta valittiin kiihtyvyyksien itseisarvojen keskiarvo, kiihtyvyyden 90. persentiili ja kiihtyvyyden keskihajonta. Itseisarvoa käytettiin, jotta negatiiviset ja positiiviset kiihtyvyydet eivät vie keskiarvoa aina lähelle nollaa.

Kulikutavantunnistusalgoritmistä tehtiin ensin versio, jossa käytettiin pelkkiä yksinkertaisia päättelysääntöjä. Päättelysäännöt käytännössä valitsevat yhden kulkutavan jonkin tai joidenkin ominaisuuksien perustella ja sulkevat muut vaihtoehdot pois. Yksinkertaiset päättelysäännöt toimivat kohtuullisen hyvin silloin, kun kulkutavalla on jokin vain sille ominainen ominaisuus, esimerkiksi kävelyllä alhainen keskinopeus. Monimutkaistamalla päättelysääntöjä voidaan ottaa huomioon useampia ominaisuuksia, mutta kovin monimutkaisista säännöistä on vaikea seurata, miten päättely etenee.

Kun halutaan ottaa huomioon useita kulkutapojen ominaisuuksia, kannattaa ottaa käyttöön pelkkiä päättelysääntöjä kehittyneempi järjestelmä. Sumea logiikka Schuesslerin ja Axhausenin (2008) tapaan on yksi vaihtoehto, mutta siihen ei tässä työssä ryhdytty, koska sumean logiikan toteuttaminen on oma projektinsa²². Sen sijaan kokeiltiin toisenlaista lähestymistapaa, jossa edellisessä vaiheessa muodostuneet monimutkaiset

²¹ Voidaan helposti todistaa, että matkan keskinopeus saadaan suoraan nopeuksien keskiarvosta silloin, kun joka välissä on kuljettu yhtä pitkä aika, tässä tapauksessa 1 sekunti. (Ongelmia aiheutuu niistä kohdista, joissa havaintoja ei ole tallentunut sekunnin välein, mutta matkan laskeminen koordinaateista ja sen jakaminen käytetyllä ajalla ei myöskään olisi täydellinen ratkaisu, koska koordinaattien epätarkkuuden takia matka olisi liian pitkä, ja sitä pitäisi korjata.)

²² Sumean logiikan käyttämiseen on myös valmiita ohjelmia tarjolla, esimerkiksi Schuessler ja Axhausen (2008) käyttivät vapaan lähdekoodin ohjelmaa.

päätelysäännöt muotoiltiin ymmärrettävämmäksi pisteytysjärjestelmäksi, jossa kulkutavoille jaettiin pisteitä sen mukaan, mihin kulkutapoihin eri ominaisuudet viittasivat.

Aineistonjakamisalgoritmi

Aineistonjakamisalgoritmissa huomioon otettavia asioita ovat pysähdysten pituudet, siis se miten pitkiä pysähdyksiä erilaiset toiminnot aiheuttavat, ja toisaalta, miten pitkiä pysähdyksiä liikenne itsessään aiheuttaa. Pysähdysten pituudet kuitenkin vaihtelevat riippuen tilanteesta, alueesta, ruuhkaisuudesta, kellonajasta ja liikennevälineestä. Yksiselitteisten parametrien määrittelemine on mahdotonta, toisin sanoen kaikki valinnat antavat johonkin suuntaan vääriä tuloksia.

Koska aineistonjakamisalgoritmi ei ollut työn pääasia, sitä tehtiin lähinnä kokeilemalla, mikä voisi toimia, eikä tutkimalla, mikä olisi teoreettisesti oikea vaihtoehto. Käytännössä siis tehtiin pääpiirteittäin toimiva versio, kokeiltiin pienellä määrällä aineistoa, miten hyvin se toimii, ja korjattiin pahimpia ongelmia säätämällä parametreja tai lisäämällä päätelysääntöjä.

Koska ratkaistavana oli myös paikallaan olevan GPS-laitteen tallentamien pisteiden tunnistaminen ja poistaminen, päätettiin yrittää tehdä se aineiston pilkkomisen yhteydessä. Paikallaan olevan laitteen tallentama ”reitti” poukkoilee edestakaisin ja siinä on lyhyitä pysähdyksiä muistuttavia kohtia, joissa nopeus on alhainen, joten lähestymistavaksi otettiin etsiä kaikki pysähdykset tai pysähdykseltä vaikuttavat kohdat.

Tällaisella metodilla löydetään paljon enemmän pysähdykseltä näyttäviä kohtia kuin aitoja pysähdyksiä, ja osa pysähdyksistä on aitoja, mutta ei katkaise matkaa (esimerkiksi bussipysäkit ja risteykset). Tällöin ollaan käänteisen ongelman edessä, eli monta pientä osaa pitäisi yhdistää osiksi, joilla on käytetty vain yhtä kulkutapaa. Tietoa siitä, missä ylimääräiset pysähdykset on tehty, voidaan jossain määrin käyttää hyväksi myös kulkutavan tunnistamisessa. Tässä pysähdyskohtia on vertailtu joukkoliikenteen pysäkkeihin. Jos käytössä olisi kattavat tiedot teistä ja kaduista sekä kevyen liikenteen reiteistä, voitaisiin ehkä myös tutkia, tapahtuvatko pysähdykset kevyen liikenteen reitillä kaukana autoteistä, ilman että koko reittiä tarvitsisi sovittaa verkolle.

Yksinkertaistetusti algoritmin kehittämisessä ongelmana oli etsiä sopivat parametrit (pysähdysten kesto ja nopeus), joilla löydetään riittävä osa aidoista pysähdyksistä ja kulkutavanvaihtoista, mutta ei aiheuteta liikaa turhia katkoksia (esimerkiksi bussista pois jäämisessä ei välttämättä ole selkeää pysähtymiskohtaa, jos bussista hyppää ulos heti oven auetessa ja lähtee heti kävelemään). Kun tämä ongelma oli ratkaistu riittävän hyvin, toinen ongelma oli yhdistellä aikaansaadut osat järkeviksi kokonaisuuksiksi ja poistaa ne osat, jotka eivät kuvanneet todellista matkustamista.

4.3.2 Algoritmien toteutus ja testaus

Algoritmit toteutettiin c-kielellä, ja testausta tehtiin koko ajan rinnakkain toteutuksen kanssa. Kulkutavanpäättelyalgoritmista tehtiin useita versioita, joista testattiin, miten hyvin ne tunnistivat eri kulkutavat koeaineistosta, ja pahimpia ongelmia yritettiin korjata seuraavassa versiossa. Lopuksi päätelysäännöt muutettiin pisteytysjärjestelmäksi.

Jokaiselle versiolle laskettiin kokonaisuonnistumisprosentti eli oikein tunnistettujen matkojen osuus aineistosta. Arvaamalla kulkutapa tässä aineistossa esiintyvien kulkutapaosuuksien perusteella saataisiin oikein 34,5 % kulkutavoista²³, joten sitä

²³ 34,5 % on saatu laskemalla aineiston 981 matkasta kunkin kulkutavan osuus prosentteina ja laskemalla sitten yhteen prosenttien neliöt. Se pätee siis ainoastaan tälle aineistolle, mutta siihen voi verrata algoritmin tuloksia, koska algoritmia sovelletaan tähän aineistoon.

suurempi onnistumisprosentti kertoo, että algoritmilla päästään parempiin tuloksiin kuin arvaamalla. Pelkkä kokonaiSONnistumisprosentti ei kuitenkaan kerro kovin paljon algoritmin toimivuudesta, koska se riippuu pitkälti siitä, kuinka paljon tutkittavassa aineistossa on milläkin kulkutavalla tehtyjä matkoja. Käytetyssä aineistossa kävelymatkoja on paljon ja ne tunnistetaan varsin hyvin, mikä nostaa kokonaiSONnistumisprosentin luultavasti suhteettoman suureksi.

Algoritmikuvauksissa olevat ehdot on ohjelmoitu niin, että jos ensin oleva ehto täyttyy, sitä seuraavia ehtoja ei enää tarkisteta. Tällä on merkitystä esimerkiksi silloin, kun tarkastetaan, osuuko matkan alku tai loppu pysäkille. Jos matka alkaa sekä bussi- että raitiovaunupysäkin läheltä, niin jos ensi tarkistetaan bussipysäkit ja vasta sen jälkeen raitiovaunupysäkit, ainoastaan bussipysäkki tulee huomioiduksi päättelyssä²⁴. Ongelma on todellinen, koska raitiovaunupysäkkien ja juna- ja metroasemien vieressä on usein bussipysäkki.

Liitteessä 6 on kuvattu muutama versio päättelysääntöihin perustuvasta algoritmista ja liitteessä 7 pisteytysjärjestelmä. Seuraavassa on kuvattu lyhyesti, kuinka hyvin eri versiot tunnistavat eri kulkutavat ja kerrottu kunkin suurimmat puutteet sekä parannukset edelliseen versioon verrattuna. Aineistonjakamisolgitmista tehtiin vain yksi versio, jota parannettiin testitulosta mukaan. Se on esitetty liitteessä 8. Algoritmin eri versioiden muuttujat on kuvattu taulukossa 10.

Taulukko 10. Kulkutavanpäättelyalgoritmin eri versioissa käytetyt muuttujat

Versio nro		0.0	1.0	1.11	2.0	2.1	2.2	3.0	3.1
Käytetty tunnistustapa	päättelysäännöt	x	x	x					
	pisteytys				x	x	x	x	x
Käytetyt muuttujat	keskinopeus	x	x	x	x	x	x	x	x
	nopeuden 90. persentiili			x	x	x	x	x	x
	nopeuksien keskihajonta			x	x	x	x	x	x
	kiihtyvyyden itseisarvojen keskiarvo					x	x	x	
	kiihtyvyyden 90. persentiili					x	x	x	
	kiihtyvyyden keskihajonta					x	x	x	
	alku- ja loppupisteet lähellä pysäkkiä		x	x	x	x	x		
	matkalla tehdyt pysähdykset				x	x			

Yksinkertaisimmassa, pelkkään keskinopeuteen perustuvassa versiossa 0.0 (taulukko 11) kevyt liikenne tunnistetaan pelkän nopeuden perusteella melko hyvin oikein, kun oletetaan hitaat matkat jalan tai pyörällä tehdyksi ja poimitaan ne heti aluksi. Suurimmat ongelmat ovat auton ja bussin sekoittuminen keskinopeuden perusteella melkein

²⁴ Numeeriset rajat, esimerkiksi nopeudet, on ohjelmoitu niin, että esimerkiksi 8-20 km/h tarkoittaa vähintään 8 mutta alle 20 km/h. Tällä ei ole käytännön merkitystä, koska GPS-aineistosta lasketut tunnusluvut ovat harvoin tasalukuja, tasalukurajat itsessään ovat arvioita ja päättelyssä käytetään muitakin ehtoja.

kaikkiin muihin kulkumuotoihin. Raitiovaunu taas sekoittuu nopeuden perusteella polkupyörään, ja juna sekoittuu bussiin ja autoon.

Taulukko 11. Tulokset kulkutavanpäättelyalgoritmin versiosta 0.0. Mahdollisimman yksinkertainen, keskinopeuksiin perustuva algoritmi.

Oikeat matkat		Algoritmin löytämät kulkutavat								onnistumisprosentti
Kulkutapa	Matkan osia aineistossa	ei tunnis-tettu	kävely	pyörä	auto	bussi	raitio-vaunu	juna	metro	
kävely	529	0	485	36	0	2	4	0	2	91,7 %
pyörä	24	0	5	19	0	0	0	0	0	79,2 %
auto	151	0	5	32	16	53	15	10	20	10,6 %
bussi	146	0	4	42	3	60	33	0	4	41,1 %
raitiovaunu	75	0	3	68	0	0	4	0	0	5,3 %
juna	34	0	1	2	10	14	0	5	2	14,7 %
metro	22	0	1	0	4	8	1	0	8	36,4 %
yhteensä	981	0	504	199	33	137	57	15	36	
Oikein yhteensä: 60,9 %										

Hieman monimutkaisemmassa versiossa 1.0 (taulukko 12) joukkoliikenteen pysäkkien ottaminen mukaan algoritmiin aiheuttaa sen, että bussit tunnistetaan paremmin, mutta monia muitakin välineitä luullaan bussiksi. Raitiovaunumatkat tunnistetaan varsin hyvin. Juna tunnistetaan melko hyvin, mutta sekoitetaan vielä bussiin; ongelma aiheutuu muun muassa siitä, että useimmilla rautatieasemilla on myös bussipysäkkejä. Keskimäärin joukkoliikenne tunnistetaan kuitenkin jo paljon paremmin kuin edellisessä versiossa.

Ongelmia on edelleen. Pyörän tunnistus huononee; pyörämatkat sekoitetaan bussimatkoihin, koska bussipysäkkejä on niin paljon, että pyörämatkat usein alkavat läheltä bussipysäkkiä. Auto ei erotu vieläkään kovin hyvin muista välineistä. Metroa ei tunnisteta, mikä johtuu siitä, että metromatkat näyttävät alkavan ja loppuvan usein melko kauas metroasemista, koska signaali ei kuulu metrossa.

Taulukko 12. Tulokset kulkutavanpäättelyalgoritmin versiosta 1.0. Lisätietona joukkoliikenteen pysäkkien paikat.

Oikeat matkat		Algoritmin löytämät kulkutavat								onnistumisprosentti
Kulkutapa	Matkan osia aineistossa	ei tunnis-tettu	kävely	pyörä	auto	bussi	raitio-vaunu	juna	metro	
kävely	529	0	485	12	0	27	4	1	0	91,7 %
pyörä	24	0	5	9	0	10	0	0	0	37,5 %
auto	151	0	5	34	70	36	0	6	0	46,4 %
bussi	146	0	4	4	9	113	12	4	0	77,4 %
raitiovaunu	75	0	3	1	0	10	61	0	0	81,3 %
juna	34	0	1	0	6	2	0	25	0	73,5 %
metro	22	0	1	0	13	8	0	0	0	0,0 %
yhteensä	981	0	504	60	98	206	77	36	0	
Oikein yhteensä: 77,8 %										

Versiossa 1.11 (taulukko 13) mukaan otettiin keskinopeuden ja pysäkkien lisäksi nopeuden 90. persentiili ja keskihajonta. Erona edelliseen on, että polkupyörä saadaan erotettua moottoriajoneuvoista entistä paremmin, ja auton erottaminen bussista paranee myös. Kokonaisuus näyttää jo varsin hyvältä, tunnistusprosentti on 82,3. Metro on edelleen ongelma, koska sitä yritetään tunnistaa asemien perusteella, mikä ei toimi. Tämän version jälkeen kulkutavanpäättelyalgoritmista tehtiin vielä muutama kokeilu päättelysäännöillä, mutta ne eivät parantaneet tulosta oleellisesti

Taulukko 13. Tulokset kulkutavanpäättelyalgoritmin versiosta 1.11. Lisätietona joukkoliikenteen pysäkkien paikat, nopeuden 90. persentiili ja nopeuden keskihajonta.

Oikeat matkat		Algoritmin löytämät kulkutavat								onnistumis- prosentti
Kulkutapa	Matkan osia aineistossa	ei tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	529	0	485	1	17	21	4	1	0	91,7 %
pyörä	24	0	5	18	1	0	0	0	0	75,0 %
auto	151	0	5	1	103	36	0	6	0	68,2 %
bussi	146	0	4	0	14	115	9	4	0	78,8 %
raitiovaunu	75	0	3	1	3	7	61	0	0	81,3 %
juna	34	0	1	0	6	2	0	25	0	73,5 %
metro	22	0	1	0	13	8	0	0	0	0,0 %
yhteensä	981	0	504	21	157	189	74	36	0	
Oikein yhteensä: 82,3%										

Versioon 1.2 lisättiin tietoa matkan osan aikana tehdyistä pysähdyksistä; käytännössä jos osan keskinopeus oli alle 22 km/h, niin vaadittiin myös pysähdyksiä bussi- tai raitiovaunupysäkeillä jotta matka tulkittaisiin kyseisellä kulkutavalla tehdyksi. Auto- matkoiksi jo tulkituista matkoista taas tarkistettiin, onko matkalla pysähdytty monta kertaa bussipysäkillä, ja jos oli, niin tulos korjattiin bussimatkaksi. Tulokset parantuivat vain marginaalisesti, mutta koska parannuksia kuitenkin havaittiin, tietoa pysähdyksistä päätettiin kokeilla pisteytysjärjestelmässä.

Versiossa 1.3 kokeiltiin käyttää tietoa siitä, onko matkan tekijällä polkupyörä käytettävissä aina tai silloin tällöin. Tulokset eivät parantuneet. Aineistossa on kuitenkin niin vähän polkupyörämatkoja, että tästä kokeilusta ei kannata tehdä johtopäätöksiä.

Versiossa 1.4. yritettiin löytää metromatkoja etsimällä niitä keskinopeuden ja nopeuden 90. persentiilin avulla vielä algoritmin lopuksi. Metromatkoja löytyi, mutta muitakin kulkutapoja, erityisesti autoa, luultiin metrokseksi, ja kokonaistulos huononi. Metro- matkojen osalta tulos oli kuitenkin niin hyvä, että todettiin, että pisteytysjärjestelmää käyttämällä metron tunnistusta voidaan luultavasti parantaa.

Seuraavaksi kehiteltiin pisteytysjärjestelmä päättelysääntöjen pohjalta. Siitäkin tehtiin useampi versio. Pisteytysperusteet on kuvattu liitteessä 7. Liitteen taulukossa on kuvattu kaikki pisteytetyt muuttujat, mutta mukaan otetut muuttujat riippuvat siitä, mikä versio on kyseessä.

Pääsääntöisesti jokaisesta sopivasta nopeus- ja kiihtyvyyshuuttujan arvosta saa yhden pisteen, mutta koska kävelyn ja junan tunnistaminen on helpointa keskinopeuden perusteella, ne saavat keskinopeudesta 2–3 pistettä. Matkan alku- ja loppupään sekä pysähdysten osumisesta oikean kulkutavan pysäkeille saa 1–2 pistettä, ja jos pysäkeille ei osuta, vähennetään pisteitä, jolloin auton ja pyörän asema paranee. Bussia kohdellaan pysäkkipisteytyksessä hieman eri tavoin kuin muita joukkoliikenteen muotoja. Näin tehdään, koska bussipysäkkejä on tiheämmässä kuin juna- ja metroasemia ja laajemmalla alueella kuin raitiovaunupysäkkejä, jolloin auto ja pyörä pysähtyvät helpommin bussipysäkkien kuin muiden pysäkkien lähetyvillä.

Versiossa 2.0 (taulukko 14) käytettiin pisteytysjärjestelmää, jossa mukana olivat nopeusmuuttujat, pysäkit ja matkan aikana tehdyt pysähdykset. Kokonaistulos 84,2 % on lähes sama kuin päättelysääntöversiossa 1.11; jotkin kulkutavat tunnistetaan hieman paremmin ja jotkut hiukan huonommin. Huomattavaa on, että pisteytysjärjestelmässä metromatkoista saadaan tunnistettua yli puolet ilman että kovin monia ei-metromatkoja luullaan metromatkoiksi.

Taulukko 14. Tulokset kulkutavanpäättelyalgoritmin versiosta 2.0. Käytetty pisteytysjärjestelmää, jossa muuttujina keskinopeus, nopeuden 90. persentiili ja nopeuden keskihajonta, joukkoliikenteen pysäkit ja matkan aikana tehdyt pysähdykset.

Oikeat matkat		Algoritmin löytämät kulkutavat								
Matkan osia Kulkutapa aineistossa		ei								onnistumis- prosentti
		tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	529	0	478	4	16	19	8	3	1	90,4 %
pyörä	24	0	5	16	0	3	0	0	0	66,7 %
auto	151	0	5	2	100	34	1	9	0	66,2 %
bussi	146	0	3	0	12	131	0	0	0	89,7 %
raitiovaunu	75	0	0	0	2	15	58	0	0	77,3 %
juna	34	0	0	0	2	2	0	30	0	88,2 %
metro	22	0	1	0	6	2	0	0	13	59,1 %
yhteensä	981	0	492	22	138	206	67	42	14	
Oikein yhteensä: 84,2 %										

Pisteytysjärjestelmällä tehtiin myös muutamia muita kokeiluja (tulokset liitteessä 9). Ottamalla versioon 2.0 mukaan kiihtyvyyshuononeet kokonaistulos huononee prosenttiyksikön verran (versio 2.1), mutta voidaan vaikuttaa jonkin verran siihen, mitä kulkutapoja tunnistetaan paremmin ja mitä huonommin. Matkan aikana tehtyjen pysähdysten poistaminen versiosta 2.0 huonontaa kokonaistuloksen ainoastaan 81,8 prosenttiin (versio 2.2), joten pysähdysten tutkiminen ei ole välttämätöntä.

Lisäksi kokeiltiin versioita, joissa tietoa joukkoliikennepysäkkien paikasta ei käytetä ollenkaan, jotta selviäisi miten algoritmit voisivat toimia, jos paikkatietoa pysäkeistä ei ole käytettävissä tai se on liian epätarkkaa. Käytettäessä ainoastaan nopeusmuuttujia tulos oli 67,9 % ja bussi- ja metromatkoja ei tunnistettu ollenkaan (versio 3.0). Joitakin bussimatkoja voitaisiin luultavasti saada näkyviin muuttamalla pisteytystä, mutta parempi tapa lienee käyttää kiihtyvyystietoja, jotka lisäämällä saatiin kokonaistulokseksi 70,1 % ja bussi- ja metromatkoista tunnistettua noin 20 % (versio 3.1).

Kulkutavantunnistusalgoritmin ensimmäisten versioiden valmistuttua kävi pian ilmi, että automaattisesti (matkapäiväkirjassa ilmoitettujen kellonaikojen perusteella) matkoiksi ja matkan osiksi pilkotulla aineistolla on vaikea testata algoritmin toimintaa täydellisesti, koska kellonaikojen epätarkkuuden takia etenkin kävelymatkojen nopeudet näyttävät liian suurilta. Kävelymatkojen oikeampien alkamis- ja päättymiskohtien etsiminen olisi ollut mahdollista, mutta työlästä ja aikaa vievää, koska kävelymatkoja on niin paljon. Muut kulkutavat toimivat melko hyvin, mutta koska niin suuri osa matkoista tehdään kävellen, olisi hyvä nähdä, millaisia kävelynopeudet todellisuudessa ovat.

Ongelmaa yritettiin ratkaista kehittämällä jakamisalgoritmia, joka jakaisi koordinaattijonon oikeammin matkan osiksi, mutta törmättiin pahempaan ongelmaan: jakamisalgoritmin tulosten arvioiminen on vaikeaa. Automaattisesti pilkotusta aineistosta kulkutavantunnistuksen onnistumista oli helppo arvioida, sillä jokaisen jakson todellinen kulkutapa oli tiedossa, ja sitä pystyi automaattisesti vertaamaan analyysin tulokseen. Sen sijaan pilkottaessa aineistoa itse tehdyillä algoritmeilla pitäisi ensin arvioida pilkkomisen onnistuminen, eli etsiä (käsin) vastinpari jokaiselle algoritmin löytämälle jaksolle, ja tarkastaa onko se oikeista kohdista pilkottu. Näille – ei välttämättä oikein pilkotuille – jaksoille pitäisi sitten arvioida kulkutavantunnistuksen onnistumista, kun itse jakso ei välttämättä oikeasti edes koostu yhden kulkutavan pisteistä. Jakamisalgoritmin tuloksia ei siis ole edes yritetty arvioida muuten kuin sanallisesti.

Jakamisalgoritmissa on itsessäänkin vielä lukuisia ongelmia, koska se on vasta alustava versio. Ongelmia tulee esimerkiksi siitä, että eri kulkutavoilla pysähdykset pysäkeillä ovat eripituisia. Esimerkiksi (pitkän matkan) juna saattaa pysähtyä joillakin asemilla minuutin ja toisilla viisi. Tällaisiin ongelmiin voidaan pureutua tarkastelemalla kunkin kulkutavan omalaatuisuuksia ja muokkaamalla algoritmia jokaisen kulkutavan kohdalla tarpeellisin tavoin, mutta tällöin algoritmin läpinäkyvyys ja dokumentoitavuus kärsii, kun erilaisia aikavälejä ja etäisyyksiä kuvaavien parametrien määrä kasvaa. Algoritmin siirrettävyys muihin olosuhteisiin pienenee myös, kun pyritään tekemään kerättyyn otokseen mahdollisimman hyvin toimiva algoritmi. Toisaalta näin saadaan tietoa siitä, miten hyvin algoritmit teoriassa voisivat toimia, jos olisi mahdollista kerätä suuria otoksia, joilla algoritmeja voidaan kalibroida tiettyihin oloihin.

5 TULOKSET

5.1 Kenttätutkimus

Kerätty aineisto

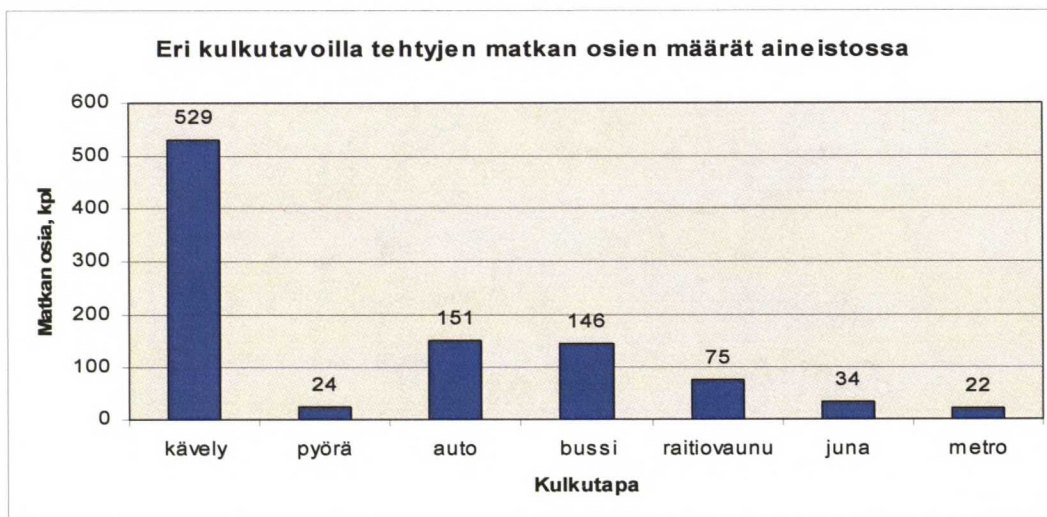
Kerätyn GPS-aineiston laatu vaihtelee sen mukaan, mitä kulkutapaa on käytetty ja kuinka avoimella alueella on liikuttu. Autoteillä aineiston laatu on pääosin hyvä; autossa, bussissa ja raitiovaunussa GPS-laite on löytänyt signaalin hyvin. Alueilla, joilla on korkeita taloja tiiviisti, voi aineiston laatu olla huonompi näilläkin kulkutavoilla. Junassa signaali hävisi melko usein. Kevyen liikenteen väylillä signaali on kuulunut hyvin, paitsi jos väylä kulkee korkeiden talojen varjossa (esimerkiksi Helsingin keskusta, jotkin lähiöt). Kevyen liikenteen alikulkutunneleissa (esimerkiksi Leppävaaran asema, Helsingin rautatieaseman ympäristö) signaali häviää. Matkojen alut puuttuvat usein GPS-laitteen käynnistymisviiveen takia.

GPS-aineisto on kuitenkin riittävän hyvää kulkutavanpäättelyalgoritmin kehittämiseen. Myöhemmin koko tutkimusaineistoa on mahdollista käyttää myös matkan tarkoituksen tai saapumispaikan tyypin päättelyyn, koska matkapäiväkirjoihin on kirjattu myös nämä tiedot vertailua varten. Osa huonommin täytetyistä matkapäiväkirjoista voidaan kuitenkin joutua jättämään tarkastelun ulkopuolelle. Voidaan myös joutua erikseen tarkastelemaan, miten hyvin kellonajan perusteella päätelty matkan päätepiste GPS-aineistossa vastaa matkapäiväkirjaan kirjattua paikkaa tai osoitetta.

Kuvassa 12 on aineistossa olevien eri kulkutavoilla tehtyjen matkan osien määrät. Suurin osa aineistossa olevista matkan osista on kävelymatkoja siksi, että jokaiseen joukkoliikennematkaan sisältyy noin kaksi kävelymatkaa, ja myös useisiin autoatkoihin sisältyy kävelyosuus. Polkupyörällä, raitiovaunulla, metrolla ja junalla tehtyjä matkoja on aineistossa varsin vähän, joten kattavampi aineisto olisi tarpeen algoritmin todellisen toimivuuden testaamiseksi. Se vaatisi kuitenkin jollain tavalla ennalta määrätyn aineiston tai niin suuren aineiston, että kaikkia kulkutapoja olisi luonnostaan riittävästi.

Tiedot osallistujista

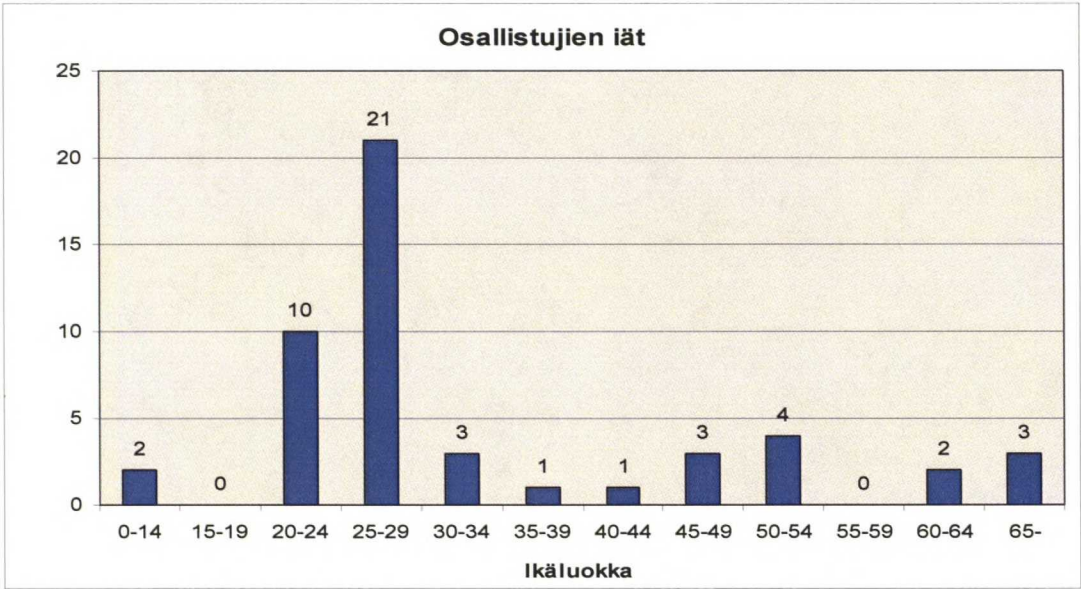
Osallistujista suurin osa oli nuoria, 20–29-vuotiaita (kuva 14). Naisia oli hieman enemmän kuin miehiä (kuva 13). Suurin osa osallistujista oli kotoisin Espoosta tai Helsingistä (kuva 15), ja suurin osa kävi töissä tai opiskelemassa Helsingissä (kuva 16).



Kuva 12. Eri kulkutavoilla tehtyt matkan osat aineistossa



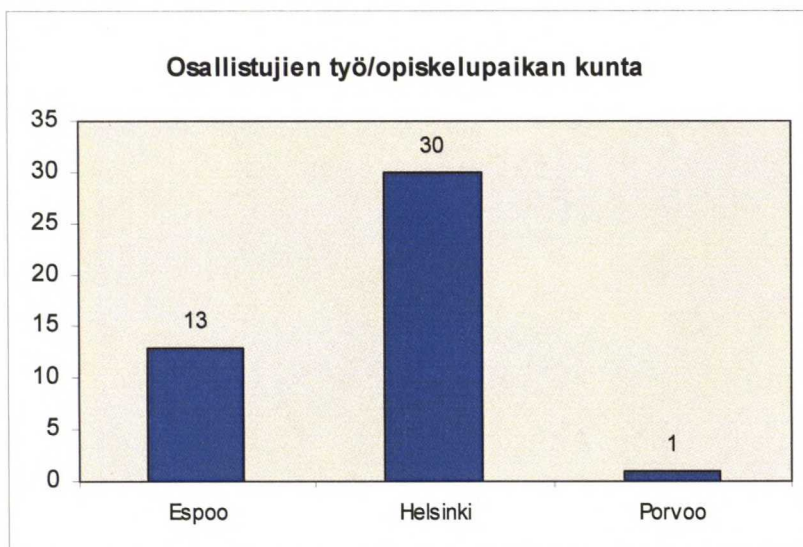
Kuva 13. Osallistujien sukupuoli



Kuva 14. Osallistujien ikä



Kuva 15. Osallistujien kotikunta



Kuva 16. Osallistujien työ-/opiskelupaikan kunta

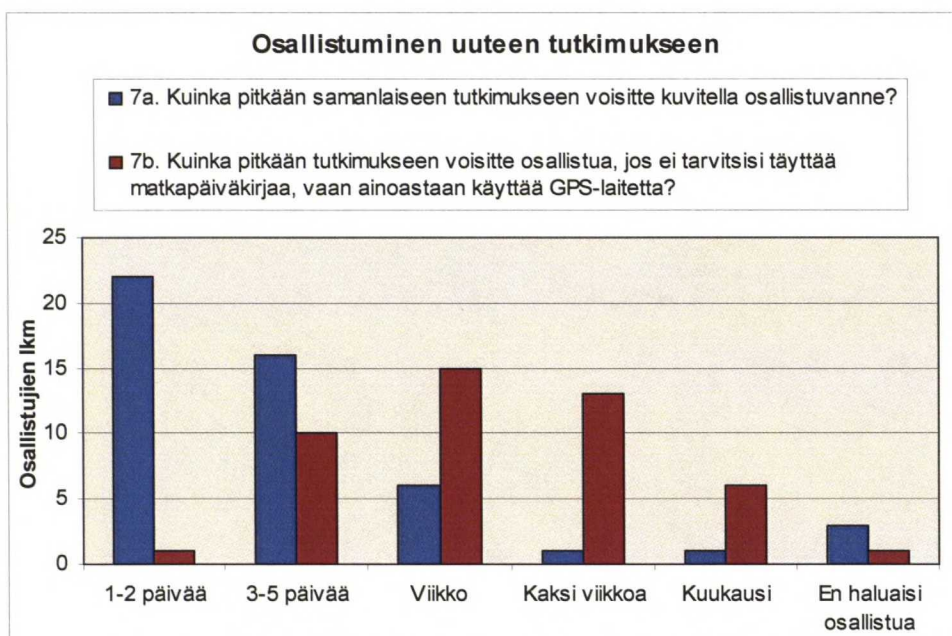
Osallistujien kokemukset

Osallistujien kokemuksia tutkimuksesta kartoitettiin kyselylomakkeella, jossa kysyttiin mielipidettä muun muassa tutkimuksen kestosta, tutkimuksen rasittavuudesta ja hankaliksi koetuista asioista. Tutkimuksen kestoa eli kahta päivää suurin osa osallistujista piti sopivana (kuva 17). Myös uuteen tutkimukseen oltiin pääosin valmiita osallistumaan (kuva 18, kysymys 7a). Jos uudessa tutkimuksessa ei tarvitsisi täyttää matkapäiväkirjaa, vaan ainoastaan pitää mukana GPS-laitetta, ihmiset olisivat tämän kyselyn mukaan valmiita osallistumaan keskimäärin pidempään tutkimukseen (kuva 18, kysymys 7b). Muutama ei haluaisi osallistua uuteen tutkimukseen. Jotkut antoivat syyksi huolen yksityisyydestä tai kiireen, mutta useat vastaajat eivät antaneet mitään erityistä syytä, miksi eivät haluaisi osallistua.

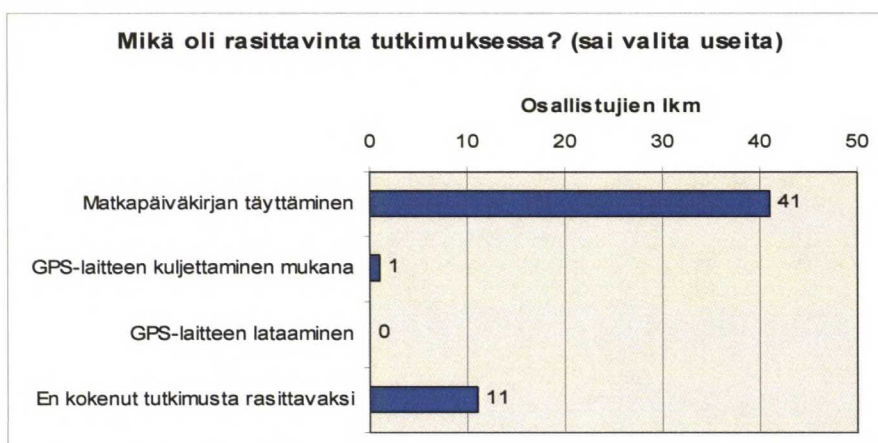
Rasittavinta tutkimuksessa oli useiden vastaajien mielestä matkapäiväkirjan täyttäminen (kuva 19). GPS-laitteen kantamista ja lataamista ei pidetty kovin rasittavana. Rasittavuus matkapäiväkirja täyttämisessä liittyi siihen, että kelloa piti tarkkailla jatkuvasti ja koko ajan piti muistaa tehdä muistiinpanoja. Eniten vaikeuksia (kuva 20) esiintyi matkapäiväkirjan täyttämisessä. Lisäksi useammalla ihmisellä oli vaikeuksia muistaa ottaa laite mukaan lähtiessään ulos. GPS-signaalin löytyminen kesti myös kauan joissakin tapauksissa, ja GPS-laitteessa olevan napin painamista matkojen alussa ja lopussa oli vaikea muistaa. Napin painamiseen liittynyt ohjeistus koettiin myös sekavaksi, ja monet painoivatkin nappia väärissä kohdissa, kuten vaihtaessaan kulkutapaa.



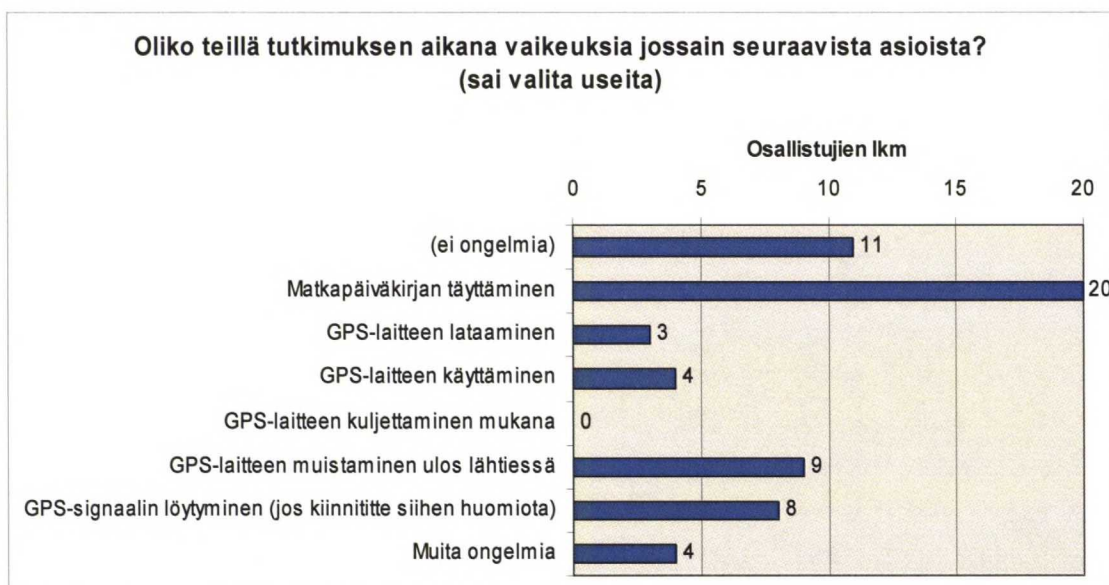
Kuva 17. Kokemukset tutkimuksen kestosta



Kuva 18. Halukkuus osallistua uuteen tutkimukseen



Kuva 19. Tutkimuksen rasittavuus



Kuva 20. Tutkimukset aikana esiintyneet vaikeudet

Matkapäiväkirjan täyttämisen useat vastaajat kokivat paitsi rasittavaksi myös vaikeaksi, mikä johtui luultavasti matkapäiväkirjan monimutkaisuudesta ja sen vaatimista tarkoista tiedoista. Matkapäiväkirjan monimutkaisuus johtui siitä, että sillä haluttiin selvittää niin monenlaista tietoa kerralla hieman tuleviakin tutkimuksia silmällä pitäen. Hyvä puoli on, että vastaavaa lomaketta, etenkin jokaisen kulkutavan vaihdon tarkkuudella, ei ole tarkoitukseen käyttänyt laajemmassa tutkimuksessa. Pitkän aikavälin tavoitteena on nimenomaan, ettei matkapäiväkirjaa tarvitsisi enää käyttää.

Matkapäiväkirjaa pidettiin myös fyysisesti liian isona, sitä oli vaikea pitää mukana. Moni osallistuja piti mukanaan vain pientä muistilappua, johon kirjasi kellonajat, ja täytti varsinaista lomaketta vasta jälkikäteen. Matkojen ja kellonaikojen kirjaaminen koettiin rasittavana ja vaivalloisena, etenkin monia osia sisältävillä joukkoliikennematkoilla ja päivinä, joihin tehtiin paljon lyhyitä matkoja. Perusteena oli muun muassa se, että matkan aikana ei välillä ole mahdollisuutta kirjoittaa mitään muistiin, ja myöhemmin on vaikea muistaa etenkin kellonaikoja ja matkan kestoa. Nämä kokemukset selittävät osaltaan, miksi etenkin lyhyitä matkoja unohdetaan perinteisissä tutkimuksissa ilmoittaa.

Muutamat vastaajat olivat myös sitä mieltä, että matkan tarkoitusta ja saapumispaikan tyyppiä kuvaavia listoja oli hankala käyttää matkapäiväkirjan takakannesta; listojen olisi pitänyt olla samalla sivulla kuin täytettävät kohdat. Matkan määritelmässä ja matkan tarkoituksen valinnassa oli jonkin verran epäselvyyttä etenkin matkaketjujen kohdalla (esimerkiksi työpaikka-kauppa-koti).

Tutkimuksessa esiintyneitä ongelmia

Osa ihmisistä, joita yritettiin rekrytoida tutkimuksen osallistujiksi, ei suostunut. Jotkut eivät halunneet, jaksaneet tai ehtineet, ja muutamat olisivat halunneet rahapalkkion vastineeksi osallistumisesta.

Tutkimuksen suorittamisessa loppuun asti ilmeni jonkin verran ongelmia. Ainakin kaksi vastaaja unohti ottaa GPS-laitteen mukaansa aamulla kotoa lähtiessään. Tässä tutkimuksessa siitä ei aiheutunut muita ongelmia kuin tutkimuksen viivästyminen, koska tutkimus siirtyi seuraaville päiville. Oikeassa tutkimuksessa, jossa tutkimuspäivät on määritetty etukäteen, laitteen unohtaminen aamulla johtaisi luultavasti siihen, että koko tutkimuspäivä olisi hyödytön. Vastaajia olisi ehkä pitänyt myös kehottaa pitämään GPS-laitetta lähellä avaimia, matkapuhelinta tai vastaavia esineitä, jotta he olisivat muistaneet ottaa sen mukaan, kuten Bricka (2008) ehdotti.

Matkapäiväkirjat ja GPS-aineistot eivät täysin vastanneet toisiaan. Yksi vastaaja ei täyttänyt matkapäiväkirjaa lainkaan, syynä oli kiire. Ainakin yhden kerran GPS-laite meni vahingossa itsestään pois päältä. Pari kertaa kävi niin, että osallistuja oli matkan aikana kirjannut ylös kellonajat ja kulkutavat, ja ajatellut täydentää loput tiedot myöhemmin, mutta ei muistanutkaan tehdä sitä. Muutaman kerran kävi myös niin, että vastaaja oli unohtanut laitteen päälle ensimmäisen tutkimuspäivän jälkeiseksi yöksi. Tällöin seuraavan päivän iltana tehdyt matkat eivät enää mahtuneet laitteen muistiin.

Kaikki osallistujat eivät olleet täysin ymmärtäneet tutkimuksen tarkoitusta, etenkin niissä tapauksissa, joissa laitteen jakelu hoidettiin välikäden kautta. Tämä johtui ehkä lähinnä siitä, että koska tutkimuslomakkeiden ja laitteiden jakelu hoidettiin pääosin henkilökohtaisesti, materiaalien joukossa ei ollut saatekirjettä, jossa olisi kerrottu tarkemmin, mistä oli kysymys. Tämä ei kuitenkaan välttämättä vaikuttanut motivaatioon täyttää matkapäiväkirjaa kunnolla, vaan muutenkin huolelliset ihmiset olivat taipuvaisempia täyttämään matkapäiväkirjaa huolellisemmin.

Tutkimuksen tavoiteaikataulussa ei pysytty. Alku sujui hyvin, kun osallistujina oli työkavereita, jolloin laitteen toimittaminen ja palauttaminen oli helppoa. Muiden

osallistujien kohdalla toimitus- ja palautusajat venyivät pidemmiksi, ja myös osallistujien vapaus valita tutkimuspäivät pidensi tutkimusaikoja hieman. Osa myös yksinkertaisesti unohti tutkimusmateriaalit tutkimuksen tehtyään, jolloin heiltä piti kysellä, ovatko he jo suorittaneet tutkimuksen. Jonkinlainen muistutus siitä, että tutkimuspäivä on tulossa, olisi varmasti hyödyllinen. Samoin tutkimusjakson jälkeen osallistujia pitäisi muistuttaa siitä, että laite pitää palauttaa.

Kustannukset ja resurssit

Kenttätutkimukseen käytettyjä resursseja voidaan arvioida karkealla tasolla. Osa kustannuksista riippuu laitteiden määrästä, osa taas osallistujien määrästä. Lisäksi kiinteänä kustannuksena on tutkimusjärjestelyjen ja lomakkeiden suunnittelu.

Laitteita oli kymmenen, ja koska ne olivat uusia, niihin piti aluksi asettaa haluttu tallennusväli ja tarkastaa, että ne toimivat. Tähän kului 10–15 minuuttia laitetta kohti eli yhteensä noin 2 tuntia. Lisäksi on tietysti otettava huomioon laitteiden hinta. Tässä tutkimuksessa käytetyt laitteet maksoivat 65 euroa, mutta lisäksi niihin oli ostettava 7 euron hintainen verkkolaturi.

Aineiston purkaminen GPS-laitteelta ja sen tallentaminen kesti noin 5 minuuttia, ja laitteen tyhjentäminen ja päivittäminen seuraavaan käyttäjää varten noin 2 minuuttia. Ajoissa on jonkin verran huomioitu sitä, että aineiston purkaminen ei aina sujunut täysin moitteettomasti, vaan välillä se kesti tavallista pidempään. Purkamiseen tarvittava aika riippuu kuitenkin käytössä olevasta GPS-laitteesta, ohjelmasta ja tietokoneesta, joten annetut ajat ovat vain suuntaa antavia. Paperilomakkeiden (taustatiedot ja loppukysely, mutta ei matkapäiväkirja) kirjaaminen sähköiseen muotoon kesti noin 5–10 minuuttia riippuen vapaan palautteen määrästä.

GPS-laitteen toimittamiseen, osallistujan opastamiseen ja laitteen noutamiseen takaisin käytettiin tässä tutkimuksessa noin 45 minuuttia osallistujaa kohti, joten jokaista osallistujaa kohden kului yhteensä noin tunti työaikaa. Tätä voitaisiin ehkä helpoiten pienentää sillä, että henkilökohtaisten tapaamisten sijaan laitteet ja lomakkeet lähetettäisiin ja palautettaisiin postitse. Toisaalta myös suurempi osallistujamäärä voisi laskea osallistujakohtaisia kustannuksia, jos toimituksia voitaisiin suunnitella tehokkaammin, ja joitakin työvaiheita voitaisiin vielä automatisoida.

5.2 Algoritminkehitys

Aineistonjakamisalgoritmi on pelkkä alustava kokeiluversio, joten siitä ei ole mielekästä antaa tuloksia.

Parhaat kehityt kulkutavanpäättelyalgoritmit tunnistivat kulkutavan oikein hieman yli 80 %:ssa kaikista matkan osista tässä tutkimuksessa käytetystä aineistosta. Jonkin algoritmin kutsuminen parhaaksi on hieman harhaanjohtavaa tässä vaiheessa, koska eri algoritmit voivat toimia eri oloissa eri tavoin. Koska erot algoritmin eri versioiden välillä ovat pieniä, toisella aineistolla voisi käydä niin, että jokin toinen algoritmi olisi parempi. Voi myös olla, että hyvin suurella aineistolla kannattaisi analysoinnin nopeuttamiseksi käyttää algoritmia, jossa on vähemmän muuttujia, mutta joka siitä huolimatta on lähes yhtä hyvä.

Jos pitää valita päättelysääntöjen ja pisteytysjärjestelmän välillä, valinta kannattaa tehdä sen mukaan, mitä kulkutapoja pitää tunnistaa ja mitä lähtötietoja on saatavilla. Algoritmit toimivat parhaimmillaan yhtä hyvin, mutta pisteytysjärjestelmällä metromatkat saadaan kuitenkin hieman paremmin tunnistettua. Toteutusvaiheessa havaittiin myös, että pisteytysjärjestelmällä on myös helpompi säädellä, minkä kulkutapojen tunnistaminen asetetaan etusijalle.

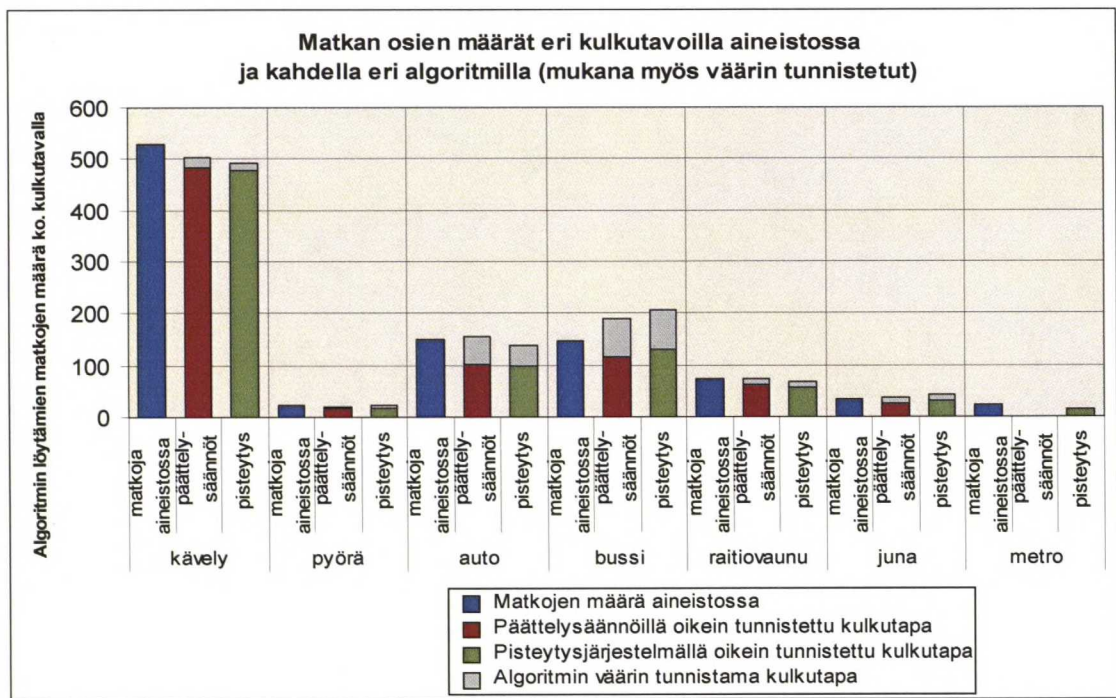
Tähän on valittu vertailtavaksi se päättelysääntöihin perustuva algoritmi, jonka kokonaistunnistumisprosentti oli suurin, ja se pisteytykseen perustuva algoritmin, jonka tunnistumisprosentti oli suurin. Tarkemmat kulkutapakohtaiset tunnistustulokset on esitetty taulukossa 15.

Taulukko 15. Vertailu päättelysääntöjen ja pisteytysjärjestelmän algoritmien toimivuudesta.

kulkutapa	matkan osia aineistossa	päättelysäännöt, versio 1.11		pisteytys, versio 2.0	
		tunnistettuja matkan osia	tunnistusprosentti	tunnistettuja matkan osia	tunnistusprosentti
kävely	529	485	91,7 %	478	90,4 %
pyörä	24	18	75,0 %	16	66,7 %
auto	151	103	68,2 %	100	66,2 %
bussi	146	115	78,8 %	131	89,7 %
raitiovaunu	75	61	81,3 %	58	77,3 %
juna	34	25	73,5 %	30	88,2 %
metro	22	0	0,0 %	13	59,1 %
yhteensä	981	807	82,3 %	826	84,2 %

Taulukosta nähdään, että pääkaupunkiseudun olosuhteissa on mahdollista tunnistaa yli 90 % kävelymatkoista, yli 80 % raitiovaunu-, juna- ja bussimatkoista ja yli 70 % pyöräilymatkoista. Automatkojen tunnistaminen näyttäisi jäävän alle 70 %:n ja metron vielä huonommaksi.

Kun taulukossa 15 tarkoituksena on ollut vertailla algoritmeja toisiinsa esittämällä vain oikein tunnistetut matkat, niin kuvassa 21 on mukana myös ne matkan osat, joiden kulkutapa tulkitaan väärin. Näin saadaan kokonaiskuva siitä, mitä tapahtuu algoritmia käytettäessä.



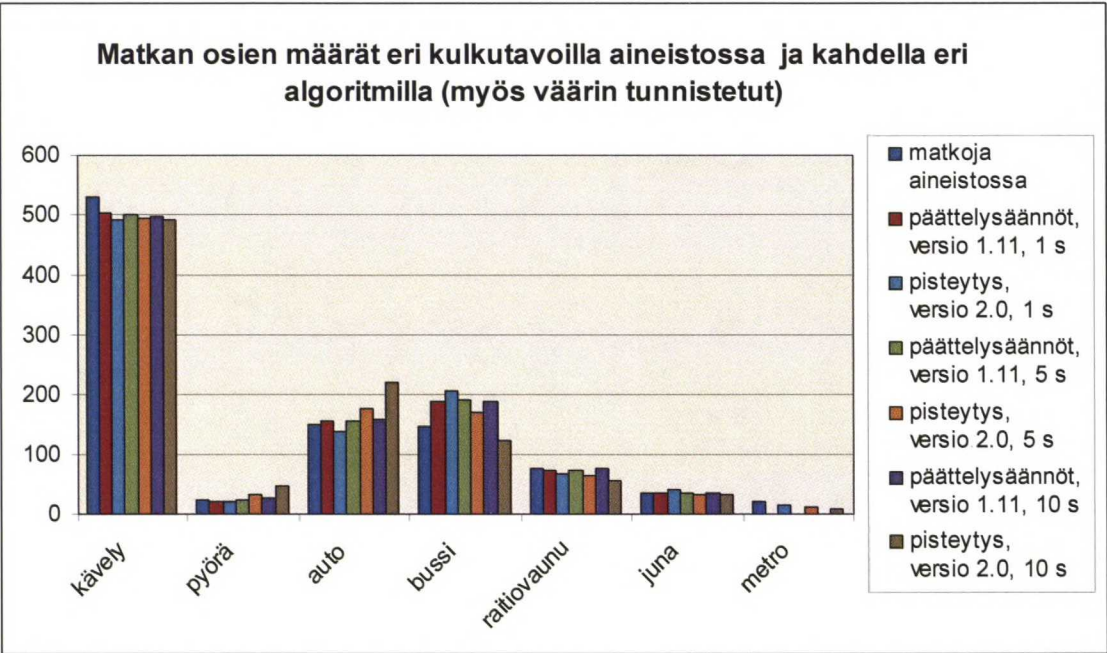
Kuva 21. Matkan osien määrät eri kulkutavoilla aineistossa (sininen) sekä kahden algoritmin löytämät kulkutavat. Vasemmanpuoleinen sininen pylväs kuvaa kävelymatkojen määrää aineistossa. Heti sen oikealla puolella oleva puna-harmaa pylväs kuvaa päättelysääntöalgoritmin (versio 1.11) löytämiä kävelymatkoja. Pylvään punainen osuus kuvaa oikein tunnistettuja kävelymatkoja ja harmaa osuus muita kuin kävelymatkoja, jotka algoritmi on kuitenkin tulkinnut kävelymatkoiksi. Vihreä-harmaa pylväs kuvaa vastaavasti pisteytysalgoritmin (versio 2.0) löytämiä kävelymatkoja. Samoin tulkitaan muut kulkutavat.

Alkuperäisessä GPS-aineistossa koordinaatit ja nopeus on tallennettu 1 sekunnin välein. Jos kehitetty algoritmi toimii riittävän hyvin myös harvemmin tallennetulla aineistolla, voidaan kerätä useamman päivän matkoja vaikka laitteen tallennuskapasiteetti olisi kohtuullisen pieni ja aineiston käsittelyyn kuluva laskenta-aika pienenee. Aineiston-jakamisalgoritmi on tehty niin, että siinä oletetaan pisteitä olevan sekunnin välein, kun etsitään pysähtymiskohtia. Se ei siis sellaisenaan voi toimia harvemmalla aineistolla, vaan pitäisi soveltaa esimerkiksi jotain kirjallisuusosassa esiteltyä menetelmää. Kuljutavanpäättelyalgoritmi taas on käytännössä riippumaton tallennusvälistä, joten sitä testattiin jäljittelemällä aineistoja, joissa koordinaatit on tallennettu 5 tai 10 sekunnin välein.

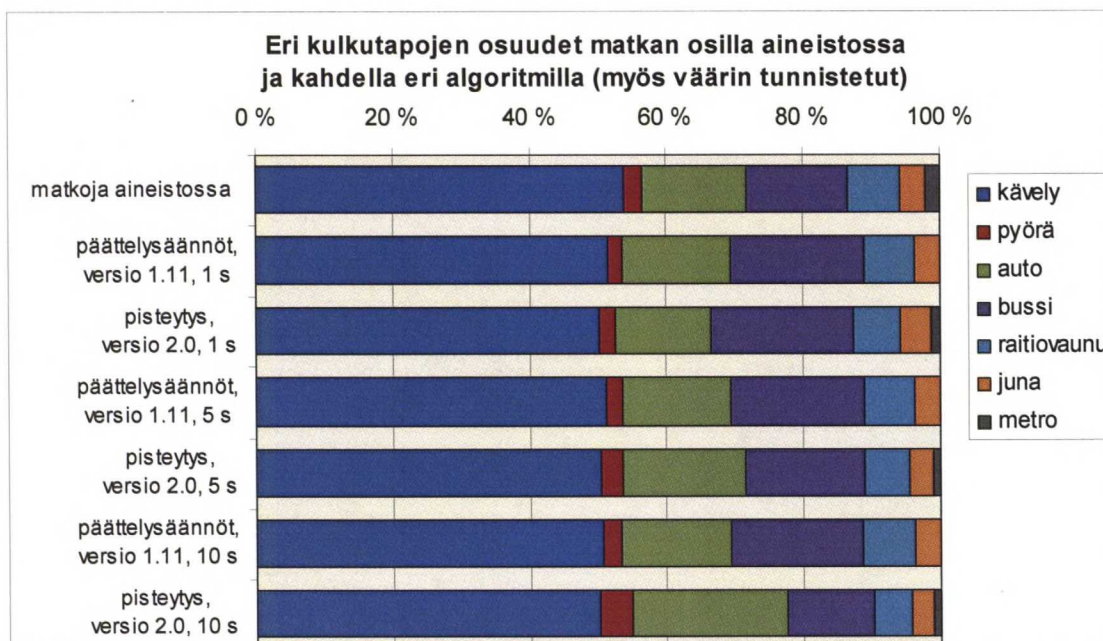
Testaus tehtiin mahdollisimman yksinkertaisesti eli poimimalla GPS-aineistosta joka 5. tai joka 10. tietue ja testaamalla näin saatuja uusia aineistoja algoritmin versioilla 1.11 ja 2.0. Tulokset ovat liitteessä 10. Havaittiin, että 5 sekunnin tallennusvälistä kokonaistunnistusprosentit heikkenivät marginaalisesti molemmilla algoritmeilla. 10 sekunnin tallennusvälistä päättelysääntöalgoritmin tulokset heikkenivät vähän lisää ja pisteytysalgoritmin tulokset selvästi. Yksittäisten kuljutapojen tunnistustarkkuus saattoi kuitenkin jopa nousta.

Kuvissa 22 ja 23 on esitetty löydettyjen matkan osien määrät ja osuudet. Samoin kuin kuvassa 21, mukana ovat myös ne matkan osat, joiden kuljutapa tulkitaan väärin, jotta saadaan kokonaiskuva algoritmin toiminnasta. Kuvassa 21 oli erotettu väärin tunnistettu osuus harmaalla, mutta kuvissa 22 ja 23 näin ei ole enää tehty, vaan kaikki algoritmin tietyllä kuljutavalla tehdyksi tulkitsemat matkat on esitetty yhtenä pylväänä.

Näyttäisi siltä, että kun väärin tunnistetut matkan osat otetaan mukaan, niin lähes kaikissa kuvien tapauksissa kävely- ja metromatkoja saadaan hieman liian vähän, pyörä-, auto-, raitiovaunu- ja junamatkoja suunnilleen oikea määrä ja bussimatkoja liikaa. Pisteytysalgoritmi 10 sekunnin välein otettu aineistoon sovellettuna on tosin poikkeus.



Kuva 22. Eri kuljutavoilla tehtyjen matkan osien määrät aineistossa ja algoritmeilla. Mukana myös väärin tunnistetut matkan osat, jotta saadaan käsitys siitä, mikä on todellinen tulos algoritmia käytettäessä. 1, 5 ja 10 s selitteessä tarkoittavat, että 1 s:n kohdalla on käytetty koko GPS-aineistoa, 5 s:n kohdalla joka viidettä aineiston tietuetta ja 10 s:n kohdalla joka kymmenettä tietuetta.



Kuva 23. Eri kulkutapojen osuudet matkan osilla aineistossa ja algoritmeilla. Mukana myös väärin tunnistetut matkan osat, jotta saadaan käsitys siitä, mikä on todellinen tulos algoritmia käytettäessä. 1, 5 ja 10 s selitteessä tarkoittavat, että 1 s:n kohdalla on käytetty koko GPS-aineistoa, 5 s:n kohdalla joka viidettä aineiston tietuetta ja 10 s:n kohdalla joka kymmenettä tietuetta.

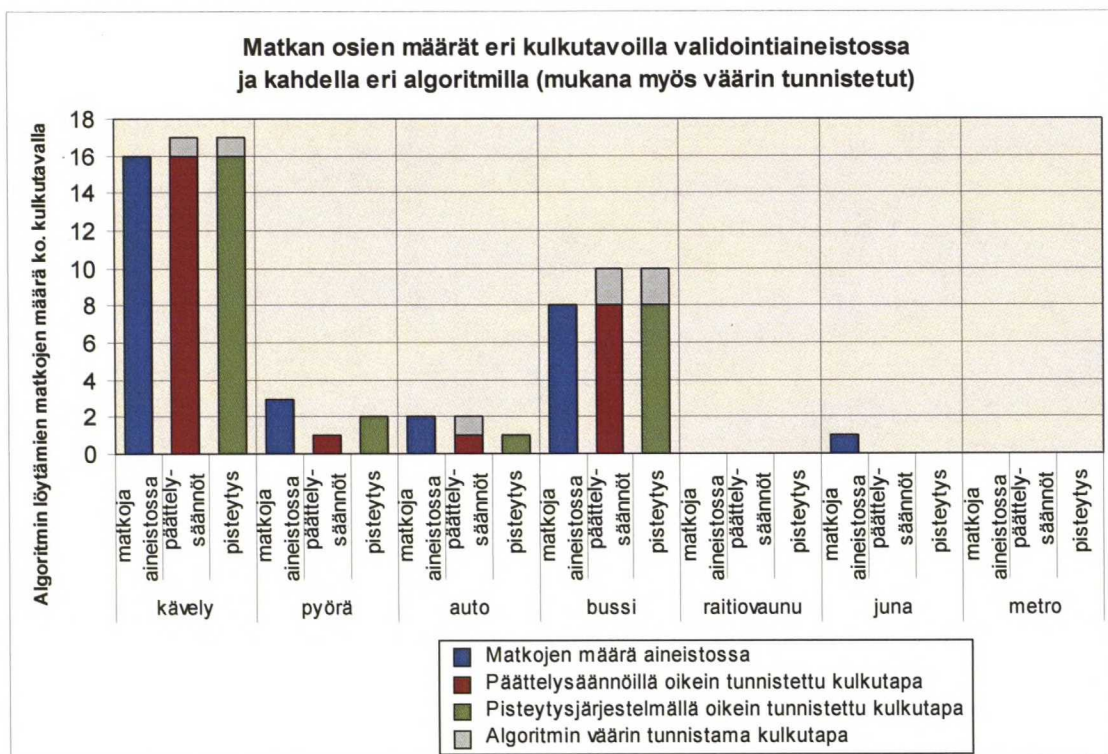
Algoritmin kehittämiseen käytetyllä GPS-aineistolla testaaminen ei anna oikeaa kuvaa algoritmin toiminnasta. Siksi algoritmeja testattiin myös pienellä aineistolla, jota ei ollut käytetty algoritmin laatimiseen. Validointiaineisto oli kuitenkin hyvin suppea. Siinä oli ainoastaan 30 matkan osaa eikä lainkaan raitiovaunu- tai metromatkoja. Validointiaineisto analysoitiin algoritmin versioilla 1.11 ja 2.0. Tulokset on esitetty liitteessä 11 ja kuvassa 24.

Testissä päättelysääntöalgoritmi tunnisti kulkutavan oikein 86,7 prosentissa tapauksista ja pisteytysalgoritmi 90 prosentissa. Tulokset ovat siis samaa luokkaa kuin varsinaisessa aineistossa, mikä on lupaavaa. Vertailuaineisto oli kuitenkin niin suppea ja matkat siinä niin samantapaisia kuin varsinaisessa aineistossa, ettei tämä pieni testi poista tarvetta testata algoritmia kattavalla, riippumattomalla aineistolla myöhemmin.

Algoritmien puutteet ja rajoitukset

Algoritmi on tehty pääkaupunkiseudulta kerätystä aineistosta, joten soveltuvuutta muualle Suomeen ei voida taata. Tätä on pohdittu hieman tarkemmin kohdassa 6.

Algoritmi tunnistaa tällä hetkellä 7 eri kulkutapaa: kävelyn, pyöräilyn, auton, bussin, raitiovaunun, junan ja metron. Niistä raitiovaunu ja metro ovat oleellisia tällä hetkellä vain Helsingissä. Automatkat ovat käytännössä henkilöautomatkoja. Aineiston bussimatkat ovat suurimmaksi osaksi pääkaupunkiseudun bussilinjoilla tehtyjä; pitkän matkan linja-autojen tunnistaminen samalla algoritmilla saattaa toimia, mutta ei välttämättä kovin hyvin. Aineistossa on niin vähän pitkiä linja-automatkoja, että sen perusteella ei pysty arvioimaan, eroavatko pitkät linja-automatkat paikallisbusseista niin paljon, että algoritmin pitäisi kohdella niitä omana kulkutapanaan. Aineiston junamatkat ovat sekä pääkaupunkiseudun lähiliikennettä että pidempiä matkoja, joten pidempien junamatkojen tunnistaminen toimii luultavasti melko hyvin.



Kuva 24. Matkan osien määrät eri kulkutavoilla validointiaineistossa (sininen), sekä kahden algoritmin (päättyksääntöversio 1.11 ja pisteytysversio 2.0) löytämät kulkutavat. Tulkintaohje on sama kuin kuvassa 21. Tämä kuva perustuu siis riippumattomaan lisäaineistoon, jota ei käytetty algoritmin tekemiseen.

Pitkillä matkoilla ja lyhyillä matkoilla niin junien kuin linja-autojenkin nopeus eroaa varmasti niiden nopeuksista lyhyillä matkoilla. Kiihtyvyydet eivät luultavasti vaihtelee yhtä paljon kuin nopeudet. Pysähdyksiä tulee lyhyillä matkoilla suhteellisesti enemmän, mutta pysähdysten vaikutus algoritmin lopputulokseen riippuu tietenkin siitä, miten algoritmi huomioi pysähdykset.

Polkupyörämatkoja oli aineistossa varsin vähän, joten polkupyörän osalta algoritmin parametrit perustuvat aineiston lisäksi intuition mahdollisista ja yleisistä nopeuksista. Toisaalta kävelynopeudet näyttivät liian suurilta lyhyillä kävelymatkoilla edellä mainittujen kellonaikojen epätarkkuuksien takia, joten myös kävelyn tunnistus perustuu osittain intuition. Testaamalla algoritmia kaikki kulkutavat kattavalla laajalla aineistolla, jota ei ole käytetty algoritmin kehittämiseen, nähtäisiin kuinka hyvin arvioissa on osuttu oikeaan.

Kevyen liikenteen osalta yritetään erottaa ainoastaan kävely ja pyöräily, harvinaisempia kulkutapoja kuten rullaluistelua ei kannata yrittää huomioida erikseen. Samoin mopot ja moottoripyörät jäävät tunnistamatta, ja ne tulkitaan luultavasti autoiksi tai polkupyöriksi. Lautta- ja laivamatkoja ei ole käsitelty millään tavalla, eikä ole tietoa miksi kulkutavaksi ne tulkittaisiin. Lentoliikennettä ei myöskään ole huomioita.

Nyt tehty algoritmi osaa ainoastaan tulkita kulkutavan yhdestä matkan osasta, joka on tehty yhdellä kulkutavalla. Tosin lyhyt kävely matkan alussa tai lopussa ei välttämättä haittaisi tunnistusta. Algoritmi ei vielä yhdistele matkan osia matkoiksi.

6 POHDINTAA

Miksi pohdintaa

Työn aikana nousi muun muassa ohjausryhmän kokouksissa esiin joitain mielenkiintoisia, huomioon otettavia tai pohdiskelun arvoisia asioita, jotka eivät kuitenkaan ole varsinaisia tuloksia tai johtopäätöksiä, koska niistä ei vielä tiedetä tarpeeksi. Osaa niistä on käsitelty aiheeseen liittyvissä artikkeleissa, toisia ei. Jotkut asiat ovat kiinnostavia teorian ja toiset käytännön kannalta. Samoin algoritminkehitysvaiheessa tuli mieleen paljon asioita, joita olisi ollut mielenkiintoista testata tai tutkia, mutta ei ollut aikaa tai sopivaa aineistoa. Seuraavassa on pohdiskeltu näitä asioita.

Kulikutavanpäättelyalgoritmi

Kulikutavantunnistusalgoritmia kehitettäessä ja parametrien ja niiden arvojen valinnassa jouduttiin tilanteeseen, jossa muutokset paransivat jonkin kulikutavan tunnistamista mutta huononsivat samalla muiden tunnistamista. Jossain vaiheessa joudutaan siis ottamaan kantaa siihen, mitkä kulutavat on tärkeintä erottaa toisistaan ja tunnistaa oikein ja mitkä voidaan jättää vähemmälle huomiolle.

Näkemykset siitä, mitkä kulutavat on tärkeintä erottaa toisistaan, vaihtelevat ihmisryhmien ja tutkimustulosten käyttötarkoituksen mukaan. Kevyen liikenteen erottaminen moottoroidusta ja joukkoliikenteen erottaminen henkilöautoista on tärkeintä, mutta mitä tarkemmin halutaan tutkia ja ennustaa liikennettä, sitä tarkemmin myös eri joukkoliikennevälineet pitää pystyä erottamaan toisistaan. (Ohjausryhmä.)

Kävely on helppo tunnistaa, mutta pyörän ja auton tunnistustulokset voisivat olla parempiakin. Jos auton ja pyörän tunnistamista saadaan parannettua, voisivat tulokset olla jo käyttökelpoisia ainakin alueilla, joilla joukkoliikenne on vähäistä. Jos joukkoliikenteen vaihtoehdot ovat nopea, omalla väylällään kulkeva juna ja linja-auto, voi auton erottaminen linja-autosta osoittautua vaikeammaksi kuin junan tunnistaminen. Joukkoliikennevälineiden keskinäistä tärkeysjärjestystä joudutaan luultavasti pohtimaan silloin, kun joukkoliikennevälineitä on useampia, ja ne muistuttavat toisiaan ja autoa.

Edellä oleva liittyy myös kysymykseen algoritmin siirrettävyydestä. Voidaanko esimerkiksi yhden kaupungin alueelta kerätyn aineiston perusteella tehtyä algoritmia käyttää toisessa kaupungissa tai aivan toisessa maassa? Voidaanko kaupungista kerätyn aineiston perusteella tehtyä algoritmia käyttää maaseudulla? Vai pitäisikö aineiston alun perinkin olla niin kattavaa, että algoritmia voidaan käyttää erilaisilla alueilla? Jos aineisto kattaa esimerkiksi kokonaisen maan, kärsiikö algoritmi siitä, että liikenneolot ovat niin vaihtelevia? Saataisiinko parempia tuloksia, jos käytössä olisi useampia algoritmeja, ja niistä valittaisiin sopivin, tai jos parametrit vaihtelisivat sen mukaan, millaisella alueella ollaan, niin että esimerkiksi kaupunkiseuduille olisi oma versionsa?

Algoritmin siirrettävyys riippuu luultavasti siitä, mitä kulutapoja algoritmi tunnistaa, kuinka paljon turvaudutaan reittien ja pysäkkien sijaintiin, mitä kaikkia kulutapoja on tarpeen tunnistaa, mihin tuloksia käytetään ja miten samanlaisia eri joukkoliikenteen muodot ovat eri kaupungeissa ja maissa. Selvää on, että esimerkiksi Helsingin raitiovaunun nopeuden ja kiihtyvyyden perusteella tunnistava algoritmi tuskin tunnistaa kovin hyvin raitiovaunuja muissa kaupungeissa. Jos taas eri joukkoliikennevälineiden tunnistaminen saadaan sidottua reitteihin, jotka voidaan esittää yleisessä muodossa (pelkkinä koordinaatteina), niin eri välineitä voi olla mielivaltainen määrä. Jo nykyisin olemassa olevia rekistereitä ja tietokantoja Suomen joukkoliikenteen reiteistä ja aikatauluista voitaisiin luultavasti käyttää apuna joukkoliikennevälineiden erottamiseksi toisistaan.

Tämän työn kulkutavanpäättelyalgoritmit on tehty pääosin pääkaupunkiseudulta kerätyn GPS-aineiston pohjalta, vaikka osa matkoista ulottuukin pääkaupunkiseudun ulkopuolella. Algoritmien toimivuutta koko Suomen kattavalla aineistolla ei ole testattu, koska sellaista GPS-aineistoa ei ole käytettävissä. Arvaus on, että haja-asutusalueilla, joilla tehdään lähinnä kävely-, pyörä- ja automatkoja, algoritmi saattaisi toimia kohtuullisen hyvin ilman suuria muutoksia. Suurissa kaupungeissa, joiden joukkoliikenne saattaa erota pääkaupunkiseudun joukkoliikenteestä, algoritmi ei luultavasti toimisi sellaisenaan. Keräämällä kattavampaa aineistoa olisi luultavasti mahdollista muuttaa parametreja niin, että algoritmi toimisi koko Suomen alueella tyydyttävästi. Jos suurissa kaupungeissa, joissa joukkoliikennettä on merkittävä määrä, haluttaisiin tehdä erillistutkimuksia, voisi olla syytä kerätä aineistoa etenkin niiden joukkoliikenteestä jotta tiedettäisiin, toimiiko algoritmi.

Algoritmin oikein ja väärin tunnistamia matkoja voisi myös olla kiinnostavaa analysoida tarkemmin. Tutkimalla millaiset (esimerkiksi kestoltaan tai pituudeltaan) matkat algoritmi tunnistaa väärin voitaisiin saada tietoa siitä, miten luotettava algoritmi on erityyppisillä matkoilla. Tietoa voitaisiin ehkä käyttää hyväksi algoritmin jatkokehityksessä.

Algoritmia ei myöskään ole vielä testattu kunnolla sellaisella aineistolla, jota ei ole käytetty algoritmin laadinnassa. Testaus riittävän laajalla, riippumattomalla aineistolla näyttäisi, onko parametreja tarvetta korjailla. Tällaisen laajan vertailuaineiston kerääminen on kuitenkin varsin työlästä, joten sellaisen saamiseksi pitäisi olla jokin muukin tarkoitus kuin algoritmin validointi.

Matkan osista matkoiksi

Jotta GPS-aineistosta saataisiin tietoa, joka on vertailukelpoista perinteisiin tutkimuksiin nähden, pitää matkan osat vielä yhdistellä matkoiksi. Matkalle pitää lisäksi määrittää pääkulkutapa. Matkan ja pääkulkutavan määritelmät eivät ole mitenkään yksikäsitteisiä, mutta algoritmin tekoa varten sellaiset pitää tehdä. On myös mahdollista vaihtaa määritelmää sen mukaan, millaiseen aineistoon tuloksia halutaan verrata.

Ainakin Stopher (2008) mainitsee, että matkan määritelmään tulisi kiinnittää huomiota, kun perinteisistä tutkimusmuodoista siirrytään automatisoidunpiin menetelmiin. Usein on tehty niin, että jokin asia määritellään aktiviteetiksi, jos se vaikuttaa reittiin tai määränpään valintaan (esimerkiksi jonkun ottaminen kyytiin), mutta matkan varrella tapahtuvaa kirjeen postittamista tai kiosilla pistäytymistä ei huomioida. Näitä asioita on pohdittava, jotta saadaan aikaan johdonmukainen matkan määritelmä. (Stopher 2008.)

Matkan ja matkan tarkoituksen määrittämisen vaikeus kävivät ilmi myös useiden kentätutkimukseen osallistuneiden palautteesta. Jos tutkittava töistä tullessaan käy kaupassa ja menee sitten kotiin, montako matkaa on tehty ja mitkä ovat niiden tarkoitukset? Jos vastaukset riippuvat siitä, kuinka kauan kaupassakäynti on kestänyt ja onko kauppa kotimatkan varrella, miten algoritmilla voidaan selvittää asia automaattisesti?

Matkan osien yhdistäminen matkoiksi voidaan tehdä esimerkiksi niin, että liitetään yhteen ne matkan osat joiden välinen aikaväli on riittävän pieni. Jos prosessissa on mukana paikkatietoaineistoa ja matkojen tarkoitukset yritetään tunnistaa, matkan osien yhdistely kannattaa liittää siihen. Matkan katkaisevan aikavälin pituus voi riippua siitä, millaisessa paikassa pysähdys tapahtuu. Voidaan myös käyttää hyväksi tietoa tyypillisistä kulkutapayhdistelmistä (esimerkiksi kävely, pyörä, (kävely-)auto(-kävely), kävely/pyörä-joukkoliikenne-kävely(-joukkoliikenne-kävely)). Tosin etenkin matkan alussa oleva lyhyt kävely ei välttämättä tallennu GPS-laitteiden hitaan käynnistymisen takia.

Pääkulkutapa voidaan määritellä esimerkiksi niin, että pääkulkutavaksi otetaan se kulkutapa, jolla on kuljettu pisin matka. Tämä määrittely on yksinkertainen, mutta ei välttämättä yhtäpitävä kaikkien tutkimusten kanssa. Jos vastaaja on itse saanut määritellä pääkulkutavan, ei päättelyä voida toistaa algoritmilla. Toisaalta joidenkin kulkutapojen asemaa saatetaan haluta korostaa. Jos yksikin matkan osa on tehty esimerkiksi lentokoneella, junalla tai metrolla, niin pääkulkutavaksi määritellään vastaavasti lentokone, juna tai metro. Toisaalta, jos päästään eroon vertailutarpeesta aikaisempiin malleihin ja opitaan mallintamaan kokonaisia monella kulkutavalla tehtyjä matkaketjuja, voidaan pääkulkutavan käsitteestä luopua kokonaan.

Muuta työn aikana esille noussutta

Uusien tiedonkeruutapojen käyttöönotto tarjoaa hyvän tilaisuuden miettiä uudelleen muitakin aikojen saatossa muodostuneita tapoja ja asioita kuin edellä mainittua matkan määritelmää. Vaikka käytössä oleva tekniikka rajaa kerättävissä olevaa tietoa, pitäisi pohtia myös tulevaisuutta ja sitä mitä tietoa kerättäisiin, jos tekniikka ei olisi este. Tarvittava tekniikka voi olla jo olemassa muulla alalla, tai se voidaan kehittää lähitulevaisuudessa.

Uutta tekniikka käyttöön otettaessa on muistettava esteettömyys. Tutkittavan prototyyppi on edelleen kiusallisen usein terve suomea puhuva aikuinen (mies) – jotta myös muiden ryhmien matkat saadaan mukaan, on kiinnitettävä huomiota laitteiden käytettävyyteen, käytön helppouteen, ohjeistukseen ja vaihtoehtoisin tapoihin kerätä tiedot. Haastatteluissa on mahdollista, että joku toinen merkitsee ja ilmoittaa matkat varsinaisen tutkittavan puolesta, mutta kannettavaa laitetta on vaikeampi pitää mukana tutkittavan puolesta, ellei ole hänen seurassaan koko ajan. Lapsille laitteita ei haluta antaa eettisistä syistä, kun taas vanhemmille ihmisille ja monille muille ryhmille laitteet voivat olla vaikeita käyttää. Laitteiden olisi siis oltava hyvin helppokäyttöisiä ja mahdollisimman automaattisia.

Ideaali laite vaatisi käyttäjältä mahdollisimman vähän. Se olisi hyvin pieni ja kevyt ja muistuttaisi ehkä itsestään väreillä tai valoilla, jotta sen muistaisi pitää mukana. Laite käynnistyisi automaattisesti – esimerkiksi ajastettuna – tutkimusjakson alussa ja olisi tutkimuksen ajan virransäästötilassa paitsi silloin, kun sen kantaja on liikkeellä, jolloin se tallentaisi reitin. Laitteen akku kestäisi vähintään 2 vuorokautta, jotta sitä ei tarvitsisi ladata lyhyiden tutkimusten aikana.

Työn aikana pohdittiin myös sitä, minkä kaikkien väestöryhmien pitäisi olla pilottitutkimuksessa edustettuina, jotta kerätty GPS-aineisto olisi riittävän kattavaa. Osallistujien joukkoon saatiin sekä lapsia että eläkeläisiä, mutta tuntuma on, että aineistossa on niin paljon vaihtelua muutenkin, että eri-ikäisten aiheuttamaa vaihtelua ei tarvitse erikseen ottaa huomioon.

Eri-ikäisten ja -kuntoisten ihmisten erot näkyvät luultavasti lähinnä kävelen liikuttaessa, mutta koska kävely muutenkin erottuu muista kulkutavoista hitautensa vuoksi, keskimääräistä hitaammin käveleviä tuskin sekoitetaan muilla kulkutavoilla liikkuviin. Pikemminkin ongelmia aiheutuu keskimääräistä nopeammasta kävelystä tai lenkkeilijöistä, joita voidaan luulla pyöräilijöiksi. Hitaammin pyöräilevät taas voivat sekoittua kävelijöihin, mutta hidas pyöräily voi johtua paitsi nuoruudesta tai vanhuudesta myös ajoympäristöstä aivan kuten autoillakin; esimerkiksi keskusta-alueilla, joilla pitää väistellä jalankulkijoita, pyöräilyn nopeus on huomattavasti alhaisempi kuin alueelta toiselle siirtymiseen tarkoitetuilla pyöräteillä. Autolla ikäihmisten ehkä hieman hitaampi vauhti hukkuu erilaisten ajoympäristöjen muutenkin aiheuttamaan vaihteluun (auton nopeusalue on noin 20–140 km/h). Julkisilla kulkuvälineillä liikuttaessa taas nopeus ei riipu henkilöstä.

7 JOHTOPÄÄTÖKSET

7.1 Kirjallisuusselvitys

GPS-laitteet tarjoavat lupaavan menetelmän erilaisten liikkumistutkimusten kehittämiseen ja tarkentamiseen. GPS-laitteiden avulla on mahdollista kerätä tietoa nykyistä tarkemmin ja erilaisiin paikkatietoihin yhdistämällä selvittää asioita, joiden tutkiminen perinteisillä menetelmillä ei onnistu (esimerkiksi käytetyt reitit ja matkojen tarkat pituudet ja kestot). GPS-pohjainen tutkimus ei kuitenkaan voi täysin korvata esimerkiksi henkilöliikennetutkimuksen nykymenettelyä, koska kaikkia nykyisin kerättäviä tietoja ei pystytä keräämään GPS:llä.

Laitteiden osalta menetelmä on käyttövalmis, eli nykyisiä GPS-laitteita voidaan niiden koon, akkujen keston ja muun käytettävyyden puolesta jo käyttää tutkimuksissa ilman suurempia ongelmia. Haasteita niillä kerätyn aineiston hyödyntämiseen tuovat kuitenkin aineiston tulkitseminen ja sen laadun arvioiminen. Aineiston tulkitsemiseen tarkoitettut ohjelmistot ovat osittain vielä kehitysvaiheessa, ja tarvitaan lisää vertailua perinteisiin tutkimuksiin GPS-laitteiden avulla saatavan tiedon laadun varmistamiseksi.

GPS-laitteet ovat hyvä mahdollisuus kerätä tietoa uudella tavalla pienissä liikkumistutkimuksissa tai kokonaan uuden tyyppisissä tutkimuksissa, joissa keskitytään enemmän tutkittavien laatuun kuin määrään tai kattavuuteen, tai tutkitaan jotain muuta kuin henkilöitä, esimerkiksi matka-aikaa. Etenkin laajoissa henkilöliikennetutkimuksissa on kuitenkin kiinnitettävä erityistä huomiota tulosten laatuun ja vastausten kattavuuteen verrattuna perinteisiin tutkimuksiin.

Henkilöliikennetutkimuksissa merkittävin ongelma on ehkä satelliittien etsimisestä laitteiden käynnistymiseen aiheutuva viive. Myös huono signaalin kuuluvuus korkeiden rakennusten katveessa aiheuttaa virheitä ja puutteita aineistoon. Viivettä voidaan pienentää apumenetelmillä, ja puuttuvia osuuksia voidaan etenkin pitkillä matkoilla rekonstruoida siinä vaiheessa, kun GPS-aineistoa käsitellään ohjelmistolla. Eniten ongelmia käynnistymisviive aiheuttaakin hyvin lyhyillä matkoilla, jotka voivat huonoissa kuuluvuusoloissa jäädä kokonaan tallentumatta. Tämä on sikäli ongelmallista, että juuri lyhyiden matkojen ilmoittamatta jättäminen on ongelma myös perinteisissä tutkimuksissa.

Laajoissa tutkimuksissa tarkimpia tuloksia saadaan luultavasti ns. *prompted recall*-menetelmällä, jolla henkilöliikennetutkimuksia on toteutettu useassa maassa. Niissä vastaajat ensin kantavat GPS-laitetta mukanaan, sen jälkeen ohjelmisto analysoi tulokset, ja lopuksi vastaajille – tai osalle vastaajista, jos heitä on paljon – näytetään tulokset, ja he korjaavat virheellisesti tulkitut tiedot. Tällaisista tutkimuksista on osallistujille paljon enemmän vaivaa kuin pelkästä haastattelusta tai GPS-laitteen käytöstä, mutta etenkin alkuvaiheessa niistä saadaan paljon hyödyllistä tietoa algoritmien kalibrointiin, jolloin seuraavissa tutkimuksissa korjausvaihe voidaan jättää pois.

Vastausprosentit eri maissa tehdyistä henkilöliikennetutkimuksen tyyppisistä tutkimuksista ja piloteista ovat lähinnä suuntaa-antavia. Realistista voisi olettaa vastausprosentiksi noin 30–40 %²⁵. On muistettava, että osallistumishalukkuus riippuu monista tekijöistä ja huoli yksityisyydestä voi nousta esiin jos tutkimusta ei onnistuta

²⁵ Puhelinhaastatteluina toteutetun vuosien 2004–2005 valtakunnallisen henkilöliikennetutkimuksen vastausprosentti oli 67 % (Henkilöliikennetutkimus 2004–2005 (2006)). Vuosien 2010–2011 valtakunnallisen henkilöliikennetutkimuksen tarjouspyynnössä taas on alustavasti oletettu (edelleen puhelinhaastatteluina tehtävän) tutkimuksen vastausprosentiksi noin 50 %.

markkinoimaan oikein. Vastausprosentin ollessa pieni on tärkeää analysoida, ketkä tutkimuksiin osallistuvat ja ketkä eivät. Suositeltavaa olisi jo suunnitteluvaiheessa analysoida perinteisten tutkimusten ja muiden maiden GPS-tutkimusten vastaajien ominaisuuksia ja pohtia, miten lisätään epätodennäköisimmin osallistuvien osallistumishalukkuutta.

Menetelmä ei siis ole vielä täysin valmis, mutta koska GPS-piloteista on monissa paikoissa saatu hyviä tuloksia ja edetty suurempiin projekteihin, kannattaa seurata mitä muualla tapahtuu. Jos halua riittää, saadaan ohjelmistoja luultavasti kehitettyä niin hyviksi, että niillä saatu tieto on vertailukelpoista aikaisempiin tutkimuksiin nähden. Aina on pidettävä mielessä tutkimuksen tavoitteet ja mietittävä, kannattaako GPS:n käyttö tiettyssä tutkimuksessa; saadaanko sen avulla parempaa tietoa tai kustannussäästöjä?

7.2 Kenttätutkimus

GPS-laitteiden avulla toteutettu kenttätutkimus onnistui odotusten mukaisesti, ja siinä kerättiin GPS-aineiston lisäksi paperinen matkapäiväkirja-aineisto vertailuaineistoksi. Vertailuaineiston keräämiseen käytetty matkapäiväkirja oli kuitenkin niin monimutkainen, että sellaisen käyttämistä ei suositella jatkossa, vaan esimerkiksi tietojen tarkistaminen osallistujilta jälkikäteen toimisi luultavasti paremmin. Hyvin lyhyiden, alle pari minuuttia kestävien matkojen – esimerkiksi kävely parkkipaikalle, lähellä olevalle bussipysäkille tai lähikauppaan – löytäminen GPS-aineistosta riittävän tarkasti pelkän matkapäiväkirjaan merkityn kellonajan perusteella ei myöskään onnistunut kunnolla, koska osallistujan kellonaika saattoi poiketa GPS-laitteen kellonajasta minuutin tai pari.

GPS-laitteita, samoin kuin mitä tahansa elektronisia laitteita käytettäessä on muistettava, että riskit ovat hieman erilaisia kuin paperisiin päiväkirjoihin tai haastatteluihin perustuvissa tutkimuksissa. Laitteiden käytettävyyteen ja toimintavarmuuteen (kestävyys, akun kesto) on kiinnitettävä erityistä huomiota. Lisäksi laitehankintojen suunnittelussa on huomioitava laitteiden kosteuden- ja pakkasenkestävyys ja mahdolliset muut tekniset rajoitukset.

GPS-laitteiden kehitys on hyvin nopeaa ja laitesukupolvet seuraavat toisiaan, joten kenttätutkimuksen valmistelu on syytä aloittaa hyvissä ajoin, mutta samalla varmistaa, että sopivaksi katsottua laitetta on saatavilla riittävä määrä silloin, kun laitteiden hankkiminen tulee ajankohtaiseksi.

Osallistujille on tärkeää antaa tietoa tutkimuksen tarkoituksesta ja siitä, mihin tuloksia käytetään. Se lisää motivaatiota ja siten toivottavasti myös parantaa kerätyn tiedon laatua.

Jotta tutkimuksen suorittaminen sujuisi hyvin, on osallistujille annettava ohjeet, miten GPS-laitetta käytetään. Jos laitteet annetaan henkilökohtaisesti, voidaan samalla varmistaa, että osallistujat oppivat käyttämään laitetta, mutta jos laitteet postitetaan, pitää löytää jokin muu tapa antaa havainnolliset ohjeet. Internetissä videona oleva ohje toimii osalle väestöstä, mutta koska kaikki eivät osaa tai pysty käyttämään internetiä, voi vaikkapa kuvallinen esite olla parempi tapa.

Koska laitteen unohtaminen kotiin tutkimuspäivän aamuna on suuri riski, pitäisi osallistujia jotenkin pystyä muistuttamaan laitteen mukaan ottamisesta. Esimerkiksi tekstiviestimuistutus tutkimuspäivän aamuna tai edellisenä iltana voisi toimia; illalla lähetettävässä viestissä voitaisiin esimerkiksi kehottaa laittamaan laite näkyvälle paikalle

eteiseen. Jos tutkimuspäiviä on useita, voidaan muistuttamista laajentaa ja osallistujilta kysyä, haluavatko he vaikkapa iltaisin tekstiviestinä muistutuksen laitteen lataamisesta.

Laitteiden toimittamiseen osallistujille ja saamiseen takaisin on varattava riittävästi aikaa. Laitteen palauttamisesta tutkimuksen jälkeen kannattaa myös muistuttaa jollain tavalla. Se voidaan tehdä esimerkiksi tekstiviestillä tai sähköpostilla automaattisesti joko heti tutkimuksen päätyttyä, tai muutaman päivän päästä, jos laitetta ei ala kuulumaan takaisin.

7.3 Algoritminkehitys

Työssä kehitetty algoritmi tunnistaa etsityt kulkutavat (kävely, pyöräily, auto, bussi, juna, raitiovaunu ja metro) noin 80 %:n todennäköisyydellä, jotkut kulkutavat hiukan paremmin ja toiset huonommin. Tulos on samaa suuruusluokkaa kuin kirjallisuus-selvityksessä tutkituilla algoritmeilla saadut tulokset. Tuloksen varmistamiseksi algoritmia pitäisi kuitenkin päästä testaamaan laajalla aineistolla, jota ei ole käytetty algoritmin kehittämiseen.

Pelkällä kulkutavanpäättelyalgoritmeilla ei ole juurikaan käyttöä. Siihen on yhdistettävä vähintään virheellisten pisteiden poisto ja aineiston jakaminen matkoiksi ja matkan osiksi.

Eri kulkutavanpäättelyalgoritmien vertailu keskenään on haastavaa. Algoritmin hyvyyteen vaikuttavat sen käyttötarkoitus, käytettävissä olevat lähtötiedot ja tunnistettavat kulkutavat.

Mitä enemmän ja tarkempia lähtötietoja (esimerkiksi joukkoliikennevälineiden pysäkit ja reitit paikkatietona) on käytettävissä, sitä parempia tuloksia yleensä saadaan. Lähtötietojen tarkkuuden parantaminen ei kuitenkaan auta enää siinä vaiheessa, kun rajoittavaksi tekijäksi tulee GPS:n tarkkuus. Tarkat lähtötiedot eivät myöskään ole aina välttämättömiä, sillä joihinkin tarkoituksiin riittävän hyviä kulkutavanpäättely-algoritmeja voi olla mahdollista tehdä myös ilman mitään lähtötietoja, eli käyttäen vain GPS-aineiston tietoja.

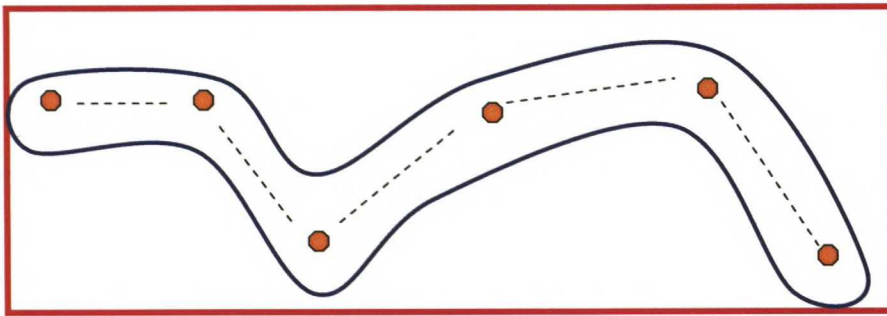
Milloin kulkutavanpäättelyalgoritmi sitten on riittävän hyvä johonkin tarkoitukseen? Yksikäsitteistä vastausta ei ole, vaan on tarpeen miettiä, mitä ja miten tarkkaa tietoa tutkimuksessa tarvitaan, ja voidaanko GPS:n avulla saada tarkempaa tietoa kuin nykyisin tai riittävän tarkkaa tietoa halvemmalla. Saavutettavissa olevien hyötyjen ja aiheutuvien kustannusten suhdetta voi yrittää arvioida, kunhan laajat tutkimukset esimerkiksi Ranskassa ja Ohiossa valmistuvat.

7.4 Kehityspolkuja

Jos henkilöliikennetutkimuksia halutaan toteuttaa GPS-avusteisina, tarvitaan kulkutavanpäättelyn lisäksi automatisointia ainakin määräpaikkojen ja niiden tyyppien sekä matkan tarkoituksen päättämisen. Muita mahdollisia tarpeita ovat reittien tunnistus (sovitus karttapohjalle), josta voidaan johtaa suorite erityyppisillä teillä, sekä matka-aikojen päättely ja ruuhkautumisen tutkiminen (jos otos on sopiva). Tarvittavia ohjelmistoja voidaan kehittää itse, tai voidaan odottaa, tuleeko markkinoille esimerkiksi paikkatieto-ohjelmistojen tai muiden valmiiden sovellusten kanssa yhdessä toimivia ohjelmistoja.

Kulkutavanpäättelyalgoritmeissa on myös edelleen parantamisen varaa. Muualla jo käytetty sumea logiikka olisi yksi vaihtoehto tässä työssä kehitettyjen algoritmien parantamiseen. Kokouksissa pohdittiin myös kulkutapojen tunnistamista niiden nopeusprofiileista hahmontunnistuksen avulla (Ohjausryhmä).

Sopivien lähtötietojen saatavuus vaikuttaa siihen, millä tavoin algoritmia voidaan kehittää eteenpäin. Väärät raitiovaunu- ja metrotunnistukset siellä, missä kyseiset välineet eivät liikennöi ollenkaan, voitaisiin poistaa pienellä vaivalla rajaamalla muutamalla koordinaatilla alue, jonka sisällä ne liikennöivät. Metro tai raitiovaunu olisi siis mahdollinen kulkutapa vain tuon rajatun alueen sisäpuolella (kuvan 25 punainen suorakulmio). Jo nyt olisi luultavasti mahdollista myös verrata koordinaatteja joukkoliikennevälineiden reitteihin, vaikkapa muodostamalla reittien ympärille käytävät (kuvan 25 sininen viiva), joihin suuren osan koordinaateista pitäisi osua, jotta sen tulkittaisiin kulkeneen kyseisellä reitillä. Siten voidaan ainakin pienentää todennäköisyyttä, että autoa luullaan joukkoliikennevälineeksi. Etenkin raitiovaunut ja metro soveltuvat tähän, koska niiden reittiverkosta Suomessa on niin lyhyt; myös junaraiteiden verkostoon voidaan verrata melko helposti, koska sekin on varsin rajallinen.



Kuva 25. Kaavakuva joukkoliikennevälineiden tunnistamisesta reittien perusteella. Oranssit pallot kuvaavat asemia tai pysäkkejä ja katkoviivat niiden välissä summittaisia reittejä. Punainen viiva kuvaa alueen, jolla kulkuvälineen käyttö ylipäättään on mahdollista (olettaen siis että kuvassa ovat kyseisen välineen kaikki pysäkit). Sinisen viivan sisään jää käytävä, johon GPS-aineiston koordinaatteja voidaan verrata, jos käytettävissä on reitin koordinaatit.

Etenkin metron tunnistaminen on hankalaa, koska GPS-signaali ei kuulu metrossa ja asemien sijainnista ei siten ole yhtä paljon hyötyä kuin muilla joukkoliikennevälineillä. Edellä mainituilla tavoilla ja tutkimalla lisäksi signaalin häviämistä metron reitillä voidaan metron tunnistamista todennäköisesti parantaa muiden joukkoliikennevälineiden tasolle. Helsingin metron tapauksessa, kun metro kulkee sekä maan alla että päällä, voitaisiin tutkia myös signaalin häviämistä kuvan 25 sinisen viivan kuvaamassa käytävässä.

Tulevaisuudessa voisi olla mahdollista liittää päättelyalgoritmeja valmiisiin liikennemalleihin, joissa on valmiina paikkatietoa kuten tiet ja kadut ja joukkoliikenteen reitit ja ehkä jopa aikataulut. Vertaamalla kerättyä GPS-aineistoa kattavaan malliin voitaisiin varsin hyvin päätellä, missä ja millä on kuljettu. Tähän tarvittava malli ei kuitenkaan olisi enää yksinkertaistava malli liikennejärjestelmästä, vaan käytännössä koko liikennejärjestelmän kuvaus. Nykyisin Suomen liikennejärjestelmän eri osia on kyllä kuvattuna jopa tällaiseen riittävällä tarkkuudella. Tietoa tuotetaan monella eri taholla, mutta sen kerääminen yhteen paikkaan on mahdollista; esimerkiksi matka.fi-palvelussa²⁶ yhdistellään tietoja eri lähteistä.

²⁶ Matka.fi -palvelu on ovelta-ovelle -reittineuvontapalvelu Suomen joukkoliikenteessä. Palvelussa yhdistellään paikallisliikenteen, VR:n junien, pitkänmatkan linja-autoliikenteen ja yhteyslauttojen reitti- ja aikataulutietoja. (Lähde <http://matka.fi>, viitattu 12.11.2009)

Henkilöliikennetutkimuksia ajatellen olennainen kysymys on myös, miten ihmiset saataisiin mukaan tutkimuksiin ja pysymään mukana niissä. Erilaiset kannustimet voisivat olla yksi keino. Rahapalkkioihin ei kuitenkaan välttämättä ole varaa, koska laajat tutkimukset ovat muutenkin kalliita. Ihmisten suhtautumisesta rahapalkkioihin ei myöskään ole varmuutta. (Ohjausryhmä.)

Karttakuva omasta liikkumisesta tutkimuksen aikana voisi olla joillekin hyvä kannustin osallistua tutkimukseen (monista kenttätutkimukseen osallistuneista oli hauska saada tutkimuksen jälkeen itselleen kartta- tai satelliittikuva, jossa näkyi heidän liikkumisensa tutkimuksen aikana). Toisaalta karttakuvan tarjoaminen palkkioksi jo etukäteen voisi myös vähentää osallistumishalukkuutta, sillä karttakuva konkretisoi sen, että osallistujien liikkumista todella seurataan tarkasti. Kannustimien tuomia hyötyjä ja haittoja pitäisikin tutkia, jotta tiedettäisiin, miten ihmiset niihin suhtautuisivat. (Ohjausryhmä.)

Henkilöliikennetutkimusten lisäksi, esimerkiksi algoritmin kehityksen rinnalla, GPS-laitteilla voidaan saada vertailutietoja eri kulkutapojen nopeuksista ja matka-ajoista erilaisissa liikenneympäristöissä. Voitaisiin esimerkiksi kulkea mahdollisimman sama reitti eri kulkuvälineillä ja vertailla pysähdyksiä ja sujuvuutta. Samalla myös algoritminkehitys hyötyisi, kun saataisiin lisätietoa kulkutapojen eroista ympäristöissä, joissa ne on vaikea erottaa toisistaan.

8 YHTEENVETO

Työn tavoitteena oli tutustua muualla maailmassa toteutettuihin kulkutavanpäättely-algoritmeihin, kerätä pilottitutkimuksessa GPS-laitteilla aineistoa pääkaupunkiseudulta ja kehittää oma kulkutavanpäättelyalgoritmi. Työssä käsiteltiin algoritmeja etenkin laajojen henkilöliikennetutkimusten näkökulmasta, mutta osa tuloksista on yleistettävissä myös muihin liikenne- ja liikkumistutkimuksiin.

Kirjallisuusselvityksessä tutustuttiin kulkutavanpäättelyalgoritmien lisäksi yleisemmin GPS-paikannukseen ja sen sovelluksiin liikkumistutkimuksissa, jo toteutettuihin ja käynnissä oleviin GPS-avusteisiin liikkumistutkimuksiin sekä algoritmeihin, joilla GPS-aineistosta voidaan automaattisesti selvittää henkilöliikennetutkimuksissa tarvittavia tietoja. Erityisesti keskityttiin kulkutavanpäättelyalgoritmeihin.

Kenttätutkimusosuudessa 50 vapaaehtoista kantoi mukanaan pientä GPS-laitetta kahden päivän ajan ja piti samalla tarkkaa päiväkirjaa matkoistaan. Menettely oli vastaajille työläs, mutta sillä saatiin selville, mitä kulkutapaa vastaaja todellisuudessa oli käyttänyt milloinkin. Vertaamalla GPS-aineistoa päiväkirjaan voitiin analysoida, miltä mikäkin kulkutapa näytti GPS-aineistossa, ja sen perusteella pohtia, miten eri kulkutavat voisi parhaiten tunnistaa.

Kirjallisuusselvityksessä vastaan tulleiden muualla toteutettujen algoritmien ja kenttätutkimuksessa kerätyn aineiston perusteella kehiteltiin kulkutavanpäättelyalgoritmi, joka toimivuuden selvittämiseksi ohjelmoitiin c-kieliseksi ohjelmaksi. Algoritmistä tehtiin erilaisia versioita, joissa käytettiin erilaisia parametreja ja lähtötietoja.

Kirjallisuusselvityksen perusteella voidaan sanoa, että GPS-paikannus liikenne- ja liikkumistutkimusten apuna on lupaava menetelmä. Mahdollisia etuja ovat vastaajien työtaakan pieneminen, tarkemmat tutkimustulokset ja tietojen tulkinnan automatisointi. Automaatio mahdollistaa työvoimavaltaisten tutkimusten toteuttamisen nykyistä halvemmalla. Riippuu tietenkin tutkimuksen tarkoituksesta, kuinka paljon hyötyä GPS-laitteiden käytöstä on. Etenkin pienemmissä tutkimuksissa GPS-laitteilla voidaan jo saada tarvittava tieto entistä helpommalla – ja jos työvoimakustannuksia pystytään vähentämään, myös halvemmalla.

Henkilöliikennetutkimuksissa ollaan käytännössä kuitenkin vielä kaukana tilanteesta, jossa paikannustieto kertoo matkan tarkoituksen ja kulkutavan yhtä tarkasti kuin haastattelu. Lisäksi monien nykyisin kerättävien tietojen, esimerkiksi mukana olevien henkilöiden määrän, auton kuljettajan ja polttoaineen kulutuksen selvittämiseen tullaan jatkossakin tarvitsemaan muita menetelmiä. Taustatietojen keräämiseenkään GPS ei tuo apua. Toisaalta tulkinta-algoritmien kehittyessä paikannuksen avulla saadaan nykyistä tarkempaa tietoa matkojen pituuksista, kestoista ja reiteistä. Myös kokonaisten matkojen unohtaminen vähenee, jos GPS-laitteita pidetään mukana koko ajan.

Kenttätutkimus alkoi GPS-laitteiden vertailulla ja hankinnalla. Tutkimukseen tarvittiin GPS-laite, joka pystyi paikantamisen lisäksi tallentamaan riittävän määrän tutkimuksessa kerättyä sijaintitietoa. Laitteiden hinnat ovat viime vuosina halventuneet, joten hinnalla ei laitteen valinnassa ollut enää kovin suurta painoarvoa. Muita kriteerejä olivat hinnan lisäksi riittävä akun kesto, helppokäyttöisyys, nopea käynnistyminen, riittävä tallennuskapasiteetti, sopiva tallennusväli ja tallennusmuoto sekä tietojen helppo siirtäminen tietokoneelle.

Kenttätutkimuksessa laadittiin osallistujille täytettävien lomakkeiden lisäksi tarkka ohjeistus, koska suurimmat riskit GPS-avusteisessa tutkimuksessa liittyivät siihen, että laitteita ei

käytettäisi niin kuin oli tarkoitus. Riskejä ovat esimerkiksi laitteen unohtaminen kotiin; laitteen päälle laittamisen unohtaminen, jos se on välissä sammutettu; lataamisen unohtaminen, jos tutkimus kestää useampia päiviä ja laitteen käyttäminen väärin, jos laitteessa on useampia toimintoja kuin päälle/pois. Tässä työssä ei olisi haitannut, jos osallistujat olisivat laitteen takia muuttaneet normaalia matkustuskäyttäytymistään tai lainanneet laitetta muille, mutta oikeissa tutkimuksissa nekin riskit on huomioitava. Lisäksi on huomioitava laitteisiin liittyvät riskit: laite voi mennä rikki kesken tutkimuksen, laite ei välttämättä löydä satelliitteja jokaisella matkalla tai akku ei kestäkään niin pitkään kuin oli ajateltu.

Kenttätutkimukseen liittyviä riskejä voidaan pienentää hyvällä ohjeistuksella laitteen käyttämisestä ja siitä, mitä ongelmatilanteissa pitää tehdä. Lisäksi on käytettävä ja kehitettävä keinoja, joilla ihmiset saadaan muistamaan tutkimus ja laitteen palauttaminen tutkimuksen jälkeen.

GPS:stä on toivottu apua muun muassa laskeviin vastausprosentteihin sitä kautta, että tutkimuksiin osallistuminen olisi helpompaa. Jotta tämä toteutuisi, on pidettävä huolta, että tutkimukseen osallistuminen todella helpottuu. Vastausprosentit toteutetuista tutkimuksista eivät indikoi korkeampaa osallistumishalukkuutta kuin perinteisissä tutkimuksissa, vaan jopa matalampaa. Ennen GPS-tutkimusta onkin syytä analysoida, mitkä ryhmät todennäköisimmin tulevat kieltäytymään tutkimukseen osallistumisesta, ja kehitettävä keinoja, joilla heidän osallistumishalukkuuttaan lisätään. Tutkimus pitäisi pystyä markkinoimaan niin, että ihmiset eivät huolestu yksityisyyden suojastaan. Jos osallistujat saadaan ymmärtämään, mitä tutkitaan ja miksi ja miten he hyötyvät siitä, kasvaa luultavasti myös motivaatio osallistua. Lisäksi on tarkkailtava, minkä tyyppiset ihmiset GPS-tutkimuksiin osallistuvat, jotta tutkimustulosten vääristymiä voidaan korjata ja tuloksia verrata aikaisempien tutkimusten tuloksiin.

Algoritminkehitysvaiheessa lähdettiin liikkeelle mahdollisimman yksinkertaisesta mallista, jossa keskinopeuden avulla pääteltiin todennäköisin kulkutapa. Algoritmeja arvioitiin sen mukaan, kuinka suurella osalla matkan osista ne tunnistivat kulkutavan oikein. Kulkutavat, joita yritettiin tunnistaa, olivat kävely, pyöräily, auto, bussi, juna, raitiovaunu ja metro.

Ensimmäisissä malleissa käytettiin yksinkertaisia päättelysääntöjä. Päättelysäännöistä tuli varsin monimutkaisia muuttujien määrän kasvaessa, joten tehtiin myös yksinkertaisempi pisteytysmalli. Siinä kulkutavoille jaettiin pisteitä jokaisesta muuttujasta, jonka arvo oli mahdollinen kyseisen kulkutavan kohdalla. Pisteytysmallilla päästiin suunnilleen samoihin onnistumislukemiin kuin päättelysäännöilläkin, mutta siinä oli helpompi hahmottaa, miten kulkutavanpäättely tapahtuu. Molemmista algoritmityypeistä tehtiin useita versioita, joissa oli mukana eri muuttujia.

Jo pelkän keskinopeuden avulla erotetaan toisistaan kävely, pyöräily ja auto melko hyvin. Ottamalla mukaan päättelysääntöalgoritmeihin joukkoliikenteen pysäkit, nopeuden 90. persentiili ja nopeuden keskihajonta tunnistettiin kulkutapa oikein yli 80 prosentissa tapauksista. Pisteytysalgoritmeissa käytettiin muuttujana myös kiihtyvyyttä, ja niissäkin tunnistettiin kulkutapa oikein yli 80 prosentissa tapauksista. Jotkin kulkutavat tunnistetaan keskimääräistä paremmin, jotkut huonommin. Esimerkiksi metroa on melko vaikea tunnistaa tässä työssä käytetyillä menetelmillä, koska GPS-signaali häviää metrotunneleissa, eivätkä matkan päät siten välttämättä osu metroasemien välittömään läheisyyteen. Toisaalta signaalin katoamista voidaan myös käyttää indikaattorina metron käyttämisestä. Myös auto on hankala, koska sillä on taipumus sekoittua melkein kaikkiin muihin kulkutapoihin.

Työssä saadut tulokset kulkutavanpäättelyalgoritmien tunnistusprosentista ovat samaa luokkaa kuin kirjallisuudessa esiintyvät. Eri kulkutavanpäättelyalgoritmien vertailu keskenään on kuitenkin haastavaa. Algoritmin hyvyyteen vaikuttavat sen käyttötarkoitus, käytettävissä olevat lähtötiedot ja tunnistettavat kulkutavat. Kävely, pyöräily ja autoilu ovat eri puolilla maailmaa usein melko samanlaisia, mutta joukkoliikennevälineet eri puolilla maailmaa ovat erilaisia. Kulkutavantunnistuksen onnistumismahdollisuudet riippuvat myös siitä, mitä lähtötietoja – joukkoliikenteen pysäkit, joukkoliikenteen reitit, tiet, kadut, kevyen liikenteen väylät – on käytettävissä.

Se, onko kulkutavanpäättelyalgoritmi riittävän hyvä tiettyyn tarkoitukseen, joudutaan arvioimaan tapauskohtaisesti. Ohjelmistojen kehittäminen vie aluksi paljon resursseja, ja hyötyä saadaan vain, jos lopulta saadaan entistä parempaa tietoa nykyistä pienemmin kustannuksin. Tässä työssä kehitettyjä algoritmeja voidaan melko helposti parantaa käyttämällä lähtötietoina täydentäviä paikkatietoja. Algoritmin rakentaminen on lopulta vain aineiston analysointia, matematiikkaa ja päättelyä. Tärkeintä on pohtia, mitä algoritmilta halutaan, mikä lopputuloksessa on tärkeintä, mitä asioita painotetaan ja mistä taas voidaan ehkä luopua.

Vaikka kulkutavanpäättelyalgoritmi olisi miten hyvä, pelkällä kulkutavanpäättelyalgoritmeilla ei ole juurikaan käyttöä, vaan siihen on yhdistettävä muita algoritmeja. Vähintään virheellisten pisteiden poisto ja aineiston jakaminen matkoiksi ja matkan osiksi on tarpeen, jotta kulkutavanpäättelyalgoritmia päästään käyttämään. Matkoiksi jakamisella päästään myös käsiksi matkojen lähtö- ja määräpaikkojen koordinaatteihin ja kellonaikoihin ja sitä kautta matkojen kestoihin. Henkilöliikennetutkimuksissa tarvitaan myös lähtö- ja määräpaikkojen tyyppin päättelyä paikkatietoaineistojen avulla. Lähtö- ja määräpaikan tyyppejä taas voidaan käyttää matkan tarkoituksen päättämiseen. Matkojen pituuksien päättämiseen tarvitaan oma algoritminsä. Se voi perustua joko suoraan GPS-aineistosta laskettuun matkan pituuteen, jota korjataan hieman lyhyemmäksi, tai se voidaan yhdistää kuljetun reitin sijoittamiseen kartalle, jolloin pituus saadaan tunnettujen tienosien pituuksista.

Aineiston keräämisen helpottuessa ja automatisoituessa ja aineistomäärien kasvaessa kannattaa myös miettiä, mikä kaikki tieto on keräämisen arvoista. Myös muista kuin liikkumistutkimuksista kertyvien GPS-aineistojen olemassaolo kannattaa muistaa, koska niihin voi kertyä hyödyllistä tietoa, jonka saamisesta tutkimuskäyttöön kannattaa neuvotella.

Muualla on jo toteutettu laajojakin GPS-avusteisia osuuksia osana perinteisiä henkilöliikennetutkimuksia, ja Ohiossa on meneillään ensimmäinen pelkkään GPS-laitteeseen perustuva henkilöliikennetutkimus. Suomessakin voitaisiin GPS-pohjaista tutkimusta kokeilla täydentämässä perinteisiä menetelmiä. Näin saataisiin vertailutietoa, joka kertoisi perinteisin menetelmin kerätyn aineiston vääristymistä. Lisäksi saataisiin tietoa GPS-tutkimuksella kerätyn aineiston puutteista.

LÄHDELUETTELO

- Abt Associates Inc. (2008). *Lehdistötiedote 23.12.2008: Abt SRBI Conducting First Fully GPS-Based Household Travel Survey in U.S.* Saatavilla online (viitattu 1.3.2009): <http://www.abtassociates.com/page.cfm?PageID=40820>
- Atmel (2009). *GPS – FAQ: Cold Start, Warm Start, and Hot Start of Antaris.* Online (viitattu 6.5.2009): http://www.atmel.com/dyn/products/faq_card.asp?faq_id=2021&family_id=665&family_name=GPS
- Bohte W., Maat K. (2008). *Deriving and Validating Trip Destinations and modes for Multi-day GPS-based Travel Surveys: a Large-scale application in the Netherlands.* Workshop paper, ISCTSC 8th International Conference on Survey methods in Transport: Harmonisation and Data Comparability (Annecy, Ranska, 25.-31.5.2008).
- Bonsall P., Wolf J., Holroyd S. (2006). *Review of the potential role of 'new technologies' in the National Travel Survey.* UK Department for Transport. Saatavilla online (viitattu 14.1.2009): <http://www.dft.gov.uk/pgr/statistics/datatablespublications/personal/methodology/ntsreports/ntsnewtechnologies.pdf>
- Bricka S. (2008) *Non-response Challenges in GPS-based Surveys.* Workshop paper, ISCTSC 8th International Conference on Survey methods in Transport: Harmonisation and Data Comparability (Annecy, Ranska 25.-31.5.2008).
- Chung E., Shalaby A. (2005) *A Trip Reconstruction Tool for GPS-based Personal Travel Surveys.* Transportation Planning and Technology, October 2005 Vol. 28, No. 5, s. 381-401.
- Flamm M., Jemelin C., Kaufmann V. (2007). *Combining person based GPS tracking and prompted recall interviews for a comprehensive investigation of travel behaviour adaptation processes during life course transitions.* Conference paper, 7th Swiss Transport Research Conference (Monte Verità / Ascona, 12.-14.9.2007).
- Gibbons G. (2008a). *China's Compass/Beidou: Back-Track or Dual Track?* Inside GNSS, Issue March/April 2008. Saatavilla online (viitattu 14.1.2009): <http://www.insidegnss.com/node/573>
- Gibbons G. (2008b). *Europe Readies Galileo Procurement - Getting Ready for an Operating Galileo System?* Inside GNSS, Issue March/April 2008. Saatavilla online (viitattu 14.1.2009): <http://www.insidegnss.com/node/585>
- Greaves S. (2004). *GIS and the Collection of Travel Survey Data.* Teoksessa Hensher D., Button K., Haynes K. ja Stopher P. (toim.), Handbook on Transport Geography and Spatial Systems (Handbook Vol. 5, Chapter 23, s. 433-449). Elsevier. Oxford.
- Henkilöliikennetutkimus 2004–2005 (2005). Internetsivut. Liikenne- ja viestintäministeriö, Tiehallinto, Ratahallintokeskus ja WSP LT-Konsultit Oy. Online, viitattu 25.2.2009. <http://hlt.fi/index.htm> (Suorat linkit on tarvittaessa annettu jokaisen viittauksen kohdalla.)
- Henkilöliikennetutkimus 2004–2005 (2006). Loppuraportti. Liikenne- ja viestintäministeriö, Tiehallinto ja Ratahallintokeskus, Helsinki. Paino Dark Oy 2006. Saatavilla online (viitattu 17.3.2009): http://www.Henkilöliikennetutkimus.fi/HTL04_loppuraportti.pdf

Henkilöliikennetutkimus 2004-2005 ks. Henkilöliikennetutkimus 2004-2005 (2005)

Jarvinen J., DeSalas J., LaMance J. (2002). *Assisted GPS: A Low-Infrastructure Approach*. GPS World, March 1, 2002. Saatavilla online (viitattu 9.11.2009): <http://www.gpsworld.com/gps/assisted-gps-a-low-infrastructure-approach-734>

Kalenoja H., Kivari M., Voltti V. (2009). *Henkilöliikennetutkimus 2010-2011. Esiselvitys*. Liikenne- ja viestintäministeriön julkaisuja 7/2009. Liikenne- ja viestintäministeriö, Helsinki. Saatavilla online (viitattu 17.3.2009): <http://www.lvm.fi/web/fi/julkaisu/view/834990>

Kitazawa K. (2006). *CAPABLE GPS report Draft 2*. Saatavilla online (viitattu 14.1.2009): http://www.casa.ucl.ac.uk/kay/others/20060707CAPABLE_GPSreport.pdf

Krygsman S., Nel J., de Jong T. (2008). *Deriving Transport Data with Cellphones: Methodological Lessons From South Africa*. Workshop paper, ISCTSC 8th International Conference on Survey methods in Transport: Harmonisation and Data Comparability (Annecy, Ranska, 25.-31.5.2008).

Landry R., Boutin P., Constantinescu A. (2006). *New anti-jamming technique for GPS and GALILEO receivers using adaptive FADP filter*. Digital Signal Processing 16 (2006) s. 255-274.

Leick A. (2004). *GPS Satellite Surveying*. (Third Edition). John Wiley & Sons, Inc., Hoboken, New Jersey.

Maanmittauslaitos (2009). Maanmittauslaitoksen internetsivut (viitattu 6.5.2009): <http://www.maanmittauslaitos.fi/GPSmittaus.htm>

Marchal P., Roux S., Yuan S., Hubert J-P., Armoogum J., Madre J-L., Lee-Gosselin M. (2008). *A Study of Non-response in the GPS Sub-sample of the French National Travel Survey 2007-2008*. Workshop paper, ISCTSC 8th International Conference on Survey methods in Transport: Harmonisation and Data Comparability (Annecy, Ranska 25.-31.5.2008).

Murakami E., Wagner D. P. (1999). *Can using global positioning system (GPS) improve trip reporting?* Transportation Research Part C 7 (1999) s. 149-165.

Ohjausryhmä. *Ohjausryhmän kokouksissa työn aikana esille tulleita ideoita*. 2008-2009.

Olsen, T. B., ja Forssell, B. (2003). *Jamming GPS. Susceptibility of Some Civil GPS Receivers*. GPS World 1.1.2003. Online (viitattu 12.5.2009): <http://www.gpsworld.com/gpsworld/article/articleDetail.jsp?id=43432>

Pasquier M., Hofmann U., Mende F., May M., Hecker D., Körner C. (2008). *Modelling and Prospects of the Audience Measurement for Outdoor Advertising Based on Data Collection Using GPS Devices (Electronic Passive Measurement System)*. Poster session paper, ISCTSC 8th International Conference on Survey methods in Transport: Harmonisation and Data Comparability (Annecy, Ranska, 25.-31.5.2008).

Stopher P. (2004). *GPS, Location, and Household Travel*. Teoksessa Hensher D., Button K., Haynes K. ja Stopher P. (toim.), *Handbook on Transport Geography and Spatial Systems* (Handbook Vol. 5, Chapter 23, s. 433-449). Elsevier. Oxford.

Stopher P. (2008). *Collecting and Processing Data from Mobile Technologies*. Resource paper for Workshop B4, ISCTSC 8th International Conference on Survey methods in Transport: Harmonisation and Data Comparability (Annecy, Ranska, 25.-31.5.2008).

Stopher P., FitzGerald C., Zhang J. (2008). *Search for a global positioning system device to measure person travel*. *Transportation Research Part C* 16 (2008) s. 350–369.

Stopher P., Greaves S. (2007). *Household travel surveys: Where are we going?* *Transportation Research Part A* 41 (2007) s. 367–38.

Särkkä S. (2007). *Modernit GPS-pohjaiset ajoneuvojärjestelmät teoriassa ja käytännössä*. Esitys Älykkään liikenteen päivässä (Espoo, Suomi, 30.10.2007). Saatavilla online (viitattu 5.5.2009): http://www.its-finland.fi/Sarkka_Indagon071030.pdf

Wolf J. (2004). *Defining GPS and Its Capabilities*. Teoksessa Hensher D., Button K., Haynes K. ja Stopher P. (toim.), *Handbook on Transport Geography and Spatial Systems* (Handbook Vol. 5, Chapter 23, s. 411-431). Elsevier. Oxford.

Wolf J., Lee M. (2008) *Synthesis of and Statistics for Recent GPS-enhanced Travel Surveys*. Workshop paper, ISCTSC 8th International Conference on Survey methods in Transport: Harmonisation and Data Comparability (Annecy, Ranska 25.-31.5.2008).

YTV ks. YTV Pääkaupunkiseudun yhteistyövaltuuskunta

YTV Pääkaupunkiseudun yhteistyövaltuuskunta (2009). *Pääkaupunkiseudun työssäkäyntialueen ajoneuvoliikenteen määräpaikkatutkimus 2008*. YTV:n julkaisu 24/2009. Valopaino Oy. Helsinki 2009. Saatavilla online (viitattu 25.10.2009): http://www.ytv.fi/NR/rdonlyres/739B137D-1FB5-45E0-8EE4-A67C789CBC02/0/loppuraportti24_2009_netti.pdf

LIITTEET

Lite 1. GPS-laitteiden vertailu

Laite	BT-335 Bluetooth Data Logger	DG-100 GPS Data Logger	iBT747 GPS Data Logger	iBT747A+ GPS Data Logger	Trip Tracker	RGM-3800 GPS Data Logger	BRT-2300 Bluetooth GPS Data Logger	RBT-1100 Bluetooth GPS	Blumax GPS-4043	Bluetooth GPS reitin-tallennuksella
Valmistaja	Globalsat	Globalsat	i-Blue	i-Blue	Amaryllo	RoyalTek	RoyalTek	RoyalTek	Blumax	Insmat
Tukeeko NMEA0183-protokollaa	Kyllä	Kyllä	Kyllä	Kyllä	Kyllä	Kyllä	Kyllä	Kyllä	Kyllä	
Heräämisaika kylmä/lämmin /kuuma	42s/38s/1s	42s/38s/1s	36s/33s/1s	35s (AGPS:llä <15s) /34s/1s	60s/38s/12s	<42s/<35s/<1s	<35s/<35s/<1s	45s/38s/8s	36s/33s/1s	
Käyttölämpötila	-20° ... 60°C	-20 ... +50 °C	-10 ... 60°C	-10 ... 60°C		-20 ... +50 °C	-20 ... +60 °C	-20 ... +60 °C	-10 ... 60°C	
Paino	75 g		kevyt	kevyt	98g	50g (+paristot)	62 g		kevyt	37 g
Mitat (LxWxH = pituus*leveys*korkeus)	72,5x40,4x26 mm	Tuumina: 3,15x2,75x0,7	72,2x46,5x20 mm	72,2x46,5x20 mm	90x58x25 mm	70x36,6x24 mm	72,2x41,0x24 mm	70x41x23 mm	72,2x46,5x20 mm	46x41,5x14m m
Paristojen / akun tyyppi	Sisään-rakennettu ladattava 1700 mAh akku	2 ladattavaa NiMh-akkua	Ladattava ja vaihdettava 1000mAh Li-ion akku	Ladattava ja vaihdettava 1000mAh Li-ion akku	Sisään-rakennettu ladattava Li-ion akku	2 AAA-paristoa / akkua	680mAh, Li-on ladattava	Vaihdettava Nokia-6100 yhteensopiva LiON-akku	Ladattava ja vaihdettava 1000mAh Li-ion akku	Sisään-rakennettu ladattava 750mAh Li-ion akku
Paristojen kesto *	25 h	jopa 24 h	25 h	25 h, jopa 30 h	jopa 32 h virransäästöllä	10 h	9 h	8 h	25 h	
Vedenkestävyys					säänkestävä	ei vedenkestävä				vesitiivis
Siru/tekniikka	SiRF Star III	SiRF Star III	MTK-piirisarja	MTK-piirisarja	SiRFstar II	SiRF Star III / SiRF GSC3f/LP	SiRF Star III		MTK-piirisarja	SiRF Star III
Kanavia	20 all-in-view-tracking	2 all-in-view-tracking	51	66		2 rinnakkaista	20 rinnakkaista	12(?)	51	

Tyhjä ruutu tarkoittaa, että tietoa ei ole löydetty

*) (luvattu, yleensä 25 °C, eli riippuu lämpötilasta, riippunee myös paristoista jos niitä voi vaihtaa)

Laite	BT-335 Bluetooth Data Logger	DG-100 GPS Data Logger	iBT747 GPS Data Logger	iBT747A+ GPS Data Logger	Trip Tracker	RGM-3800 GPS Data Logger	BRT-2300 Bluetooth GPS Data Logger	RBT-1100 Bluetooth GPS	Blumax GPS-4043	Bluetooth GPS reitin- tallennuksell a
Herkkyys	-159 dBm	-159dBm	-158 dBm	-165 dBm		-159dBm	-159dBm	-159 dBm	-158 dBm	-159 dBm (?)
Tarkkuus	Paikka: 10 m (5m jos WAAS päällä); nopeus: 0,1m/s	Paikka: 10 m (1-5m jos WAAS päällä); nopeus: 0,1m/s	Paikka: 3m	Paikka: 3m		Paikka: 10 m 90%:lle; nopeus: 0,1m/s	Paikka: 10 m 90%:lle; nopeus: 0,1m/s		Paikka: 3m	lienee sama kuin muilla vastaavan sirun laitteilla
DGPS	WAAS/ EGNOS/ MSAS	WAAS/ EGNOS/ MSAS	WAAS/ EGNOS/ MSAS	WAAS/ EGNOS/ MSAS		WAAS/ EGNOS	WAAS/ EGNOS		WAAS/ EGNOS/ MSAS	
Tallennusväli	ei kerrota, manuaalin perusteella rajaton (1s ylöspäin), oletusarvo 30s	10 s, ei voi säätää	minimi 1s, säädettävissä sekunnin tarkkuudella	minimi 1s, säädettävissä sekunnin tarkkuudella		säädettävissä 5s ... 60s, oletusarvo 15s	säädettävissä 5s ... 60s, oletusarvo 15s **		minimi 1s, säädettävissä sekunnin tarkkuudella	5s, 15s, 30s tai 1min
Mitä tallennetaan (paikka tarkoittaa tässä pituus- ja leveys- koordinaatteja)	1. Vain paikka 2. Pvm, aika, paikka, nopeus 3. Pvm, aika, paikka, korkeusasema, nopeus	Pvm, aika, paikka, korkeusasema, nopeus	Pvm, aika, paikka, korkeusasema , nopeus	Pvm, aika, paikka, korkeusasema, nopeus		A) UTC aika, paikka, korkeusasema, nopeus B) UTC aika, paikka, korkeusasema C) UTC aika, paikka	A) UTC aika, paikka, korkeusasema, nopeus; B) UTC aika, paikka, korkeusasema; C) UTC aika, paikka		Pvm, aika, paikka, korkeusasema, nopeus	
Muistin koko ja / tai montako pistettä muistiin mahtuu	16 Mb, 60 000 pistettä (40 000 jos korkeus mukana)	16 Mb, 60 000 pistettä (40 000 jos korkeus mukana)	32 Mb 150 000 pistettä	32 Mb 125 000 pistettä	8192 reittipistettä ja 20 reittiä. Paikka erilliselle muistikortille, jolle voi tallen- taa pidemmän reitin.	64 Mb A) 400 000 pistettä B) 500 000 C) 650 000	65 Mb A) 400 000 pistettä B) 500 000 C) 650 000	30 000 pistettä	32 Mb 150 000 pistettä	16 000 pistettä

Tyhjä ruutu tarkoittaa, että tietoa ei ole löydetty

**) (Internetin keskusteluryhmän mukaan myös 1s onnistuu ohjelmistopäivityksen ansiosta vaikka ei "pitäisi")

Laite	BT-335 Bluetooth Data Logger	DG-100 GPS Data Logger	iBT747 GPS Data Logger	iBT747A+ GPS Data Logger	Trip Tracker	RGM-3800 GPS Data Logger	BRT-2300 Bluetooth GPS Data Logger	RBT-1100 Bluetooth GPS	Blumax GPS-4a	Bluetooth GPS reitin- tallennuksell
Tallennusformaatit	KML, CSV, Text, RMC, GPX	KML, CSV, Text, RMC, GPX	useita	useita	NMEA	NMEA	NMEA		useita	
Voiko näyttää Google Earthissa/Mapsissa	Kyllä	Kyllä	Kyllä	Kyllä		kyllä	ilmeisesti ei		Kyllä	Kyllä
Tiedonsiirtoväylä tietokoneelle	Bluetooth	USB	USB, Bluetooth	USB, Bluetooth	USB	USB (mini)	Bluetooth 2.0	Bluetooth 1.1	USB, Bluetooth	Bluetooth 2.0 / USB
Tuleeko mukana verkkolaturi	kyllä		kyllä	ei					ei	
Hinta (arvio), ei sisällä postikuluja	95 €	117 €	65 €	69 € + verkkolaturin hintaa 9,90 €	94 £ (= 111 €)	69,95 € (/99,99 \$)	105 €	99 € / 102 €	59 € + verkkolaturin hintaa 14,90 €	74,90 €
Saatavuus Suomesta / ulkomailta	lienee tällä hetkellä saatavilla	lienee tällä hetkellä saatavilla	saatavissa Suomesta (nettikauppa, 2-5 päivää)	saatavissa Suomesta (nettikauppa, 2- 5 päivää)	Ei Suomesta luultavasti, mutta nettimyyntinä ainakin UK:sta	lienee tällä hetkellä saatavilla	Tilapäisesti loppu ko. liikkeestä, löytynee Saksasta tms.	lienee tällä hetkellä saatavilla	Saatavissa Suomesta	
Muita huomioita	Datum WGS- 84	Datum WGS- 84, Valmiiksi ohjelmoitaviss a 3 eri tallennustilaa, joita voi vaihtaa lennossa	Datum WGS- 84, Erikseen mainittu, että soveltuu myös hitais- sa kävely- nopeuksissa käytettäväksi	Korvasi mallin iBT747 helmi- kuussa 2009, uutena ominai- suutena AGPS (ratatietojen luku etukäteen internetistä)	Käytetty hollantilaisessa tutkimuksessa	Mukana karttasovellus	Mukana karttasovellus	Vaikea löytää valmistajan sivuilta	Oleellisilta osin sama laite kuin iBT747 (mm. sama ohjelmisto), mutta ilman verkkolaturia	

Tyhjä ruutu tarkoittaa, että tietoa ei ole löydetty

Laite	BT-335 Bluetooth Data Logger	DG-100 GPS Data Logger	iBT747 GPS Data Logger	iBT747A+ GPS Data Logger	Trip Tracker	RGM-3800 GPS Data Logger	BRT-2300 Bluetooth GPS Data Logger	RBT-1100 Bluetooth GPS	Blumax GPS- 4043	Bluetooth GPS reitin- tallennuksell a
Tietojen lähde (jos pvm ei mainittu niin 4.12.2008)	Valmistajan kotisivu http://www.globalsat.com.tw/eng/product_detail1_00000103.htm	Valmistajan kotisivu http://www.globalsat.com.tw/eng/product_detail1_000001090.htm	http://www.digimesta.com/index.php?page=shop.product_details&flypage=flypage_new&product_id=547&category_id=23&manufacturer_id=0&option=com_virtuemart&Itemid=63	(16.2.2009) http://www.digimes.ta.com/index.php?page=shop.product_details&flypage=flypage_new&product_id=547&category_id=23&manufacturer_id=0&option=com_virtuemart&Itemid=63&vmchk=1	http://www.amaryllo.com/almooj/outdoor-products/amaryllo-trip-tracker-gps-outdoor-device.html	Valmistajan kotisivut http://www.royaltek.com/products_dtl.php?cid=2&id=23&argPage=1&argl=3	Valmistajan kotisivut http://www.royaltek.com/products_dtl.php?cid=2&id=19&argPage=1&argl=6	http://www.digimania.fi/index.php?module=ekauppa&type=2&product=ROY-RBT-1100 , http://www.tietoa.sema.fi/eshop/index.php?sid=1&pcode=RBT-1100	(16.2.2009) http://www.eurotronic.net/products/products/gps_4043_recorder.html	http://www.akkunetti.fi/product_info.php?cPath=79&products_id=13866
Hinnan lähde (jos pvm ei mainittu niin 4.12.2008)	http://www.tietofriikki.fi/product_details.php?p=3583	http://www.tietofriikki.fi/product_details.php?p=2089	http://www.digimesta.com/index.php?page=shop.product_details&flypage=flypage_new&product_id=547&category_id=23&manufacturer_id=0&option=com_virtuemart&Itemid=63	(16.2.2009) http://www.digimes.ta.com/index.php?page=shop.product_details&flypage=flypage_new&product_id=547&category_id=23&manufacturer_id=0&option=com_virtuemart&Itemid=63&vmchk=1	http://www.pcplotter.com/price.htm	http://www.timeimage.fi/cat/product_details.php?p=9330 , http://www.semsonson.com/rorgggsdalow.htm	http://www.jt-net.fi/product_details.php?p=707	http://www.digimania.fi/index.php?module=ekauppa&type=2&product=ROY-RBT-1100 , http://www.tietoa.sema.fi/eshop/index.php?sid=1&pcode=RBT-1100	(16.2.2009) http://www.elektro.fi/?product=13194	http://www.jt-net.fi/product_details.php?p=2400
Kuva										

Tyhjä ruutu tarkoittaa, että tietoa ei ole löydetty

Liite 2. Tiedonsiirtoyhteyksien hintoja (8.10.2009)

Taulukon hinnat ovat esimerkkejä yksityisille kuluttajille tarkoitetuista liittymistä. Yrityksille on räätälöitävissä myös muita ratkaisuja. Tiedonsiirtoyhteyksien hinnat vaihtelevat nopeuden ja siirretyn datamäärän mukaan. Siihen, millä teknisellä ratkaisulla yhteys toteutetaan, ei ole otettu kantaa. Erilaiset päätelaitteet (esim. matkapuhelimet, kämmentietokoneet) tukevat erilaisia tiedonsiirtoyhteyksiä.

Operaattori	Liittymä	Hinta *	Max. nopeus **	Rajoitukset
DNA	DNA Tiedonsiirtopalvelu	1,50 €/Mt		
	DNA Matkapaketti	4,90 €/kk		sisältää 30 Mt/kk, ylittävältä osalta 1,50 €/Mt
	DNA Matkanetti	9,80 €/kk	384 kbit/s	
	DNA Matkanetti	14,80 €/kk	512 kbit/s	
	DNA Matkanetti	19,80 €/kk	1 Mbit/s	
	DNA Matkanetti	29,80 €/kk	2 Mbit/s	
Elisa	Elisa Perusdata	4,88 €/Mt		
	Elisa Datapaketti	4 €/25 Mt		Ylittävä käyttö 4 € / alkava 25 Mt
	Elisa Mobiililaajakaista 384	9,90 €/kk	384 kbit/s	
	Elisa Mobiililaajakaista 512	14,90 €/kk	512 kbit/s	
	Elisa Mobiililaajakaista 1 M	19,90 €/kk	1 Mbit/s	Edellyttää HSDPA-tekniikkaa (vähintään 1,8 Mbit/s) tukevaa päätelaitetta
	Elisa Mobiililaajakaista 2 M	29,90 €/kk	2 Mbit/s	Edellyttää HSDPA-tekniikkaa (vähintään 3,6 Mbit/s) tukevaa päätelaitetta
Saunalahti	Internet-käyttö	1,50 €/Mt		
				sisältää 25Mt/kk, ylimenevä osa liittymätyypin hinnaston mukaan (esim. 1,50 €/Mt)
	GPRS 25 -palvelu	4,00 €/kk		
	Mobiililaajakaista 384	9,80 €/kk	384 kbit/s	
	Mobiililaajakaista 512	14,80 €/kk	512 kbit/s	
	Mobiililaajakaista 1 M	19,80 €/kk	1 Mbit/s	
	Mobiililaajakaista 2 M	29,80 €/kk	2 Mbit/s	
	Mobiililaajakaista 5 M	34,90 €/kk	5 Mbit/s	
Sonera	Data Tunti	0,90 €/alkava tunti		
	Data Porras	3,90 €/kk		sisältää 20 Mt/kk, ylittävältä osalta 3,90 €/alkava 20 Mt
	Data Päivä	2,90 €/alkava 24 h		
	Data Kuukausi	14,90 €/kk		(sisältää kohtuullisen käyttö 3 Gt/kk)***
	Data Teho	29,80 €/kk	1 Mbit/s	
	Data Teho Plus	34,80 €/kk	3,6 Mbit/s	

*) Useimmissa palveluissa on lisäksi kytkentämaksu esim. 1,90 € / 2,90 € / 3,90 €

**) Yhteyden nopeus riippuu mm. alueesta jolla tiedonsiirtoyhteyttä käytetään

***) Oletettavasti muillakin operaattoreilla on olemassa jokin "kohtuullisen käytön" raja

Mt = megatavu, Gt = gigatavu

Liittymät sisältävät tiedonsiirron kotimaassa, ulkomailla tapahtuva tiedonsiirto laskutetaan erikseen.

Lähteet (8.10.2009):

<http://www.dna.fi/Yksityisille/Matkaviestinta/LiittymienPeruspalvelut/Sivut/Tiedonsiirto.aspx>

<http://www.dna.fi/Yksityisille/Matkaviestinta/LiittymienLisapalvelut/Sivut/dnaMatkapaketti.aspx>

<http://www.dna.fi/Yksityisille/Matkaviestinta/LiittymienLisapalvelut/Sivut/dnaMatkanetti384,512ja1M.aspx>

<http://www.elisa.fi/matkaviestinta/index.cfm?o=199.60>

<http://saunalahti.fi/gsm/hinnastomuut.php>

<http://saunalahti.fi/gsm/mobiililaajakaista.php>

<http://www.sonera.fi/Puhelin%20ja%20liittym%E4/Hinnasto/>

Liite 3. Tutkimuslomakkeet

GPS-tutkimus – Taustatietolomake

Taustatietolomakkeessa kysytään yleisiä henkilöliikennetutkimuksiin liittyviä kysymyksiä, sekä lisäkysymyksiä, joista saatavaa tietoa on tarkoitus käyttää GPS-aineiston tulkitsemiseen (mm. kulkutavan ja matkan tarkoituksen päättelyminen).

Vastatkaa kysymyksiin tämänhetkisen tilanteen, esimerkiksi viimeisen kuukauden, pohjalta (paikat joissa tällä hetkellä käytte usein, kulkutavat joita tällä hetkellä tai tähän vuodenaikaan käytätte).

1. Ikä

_____ vuotta

2. Sukupuoli

☐ Nainen ☐ Mies

3. Työssäkäynti: teettekö kokopäiväistä vai osa-aikaista ansiotyötä vai oletteko ansiotyössä tällä hetkellä ollenkaan (voitte valita useita vaihtoehtoja)

- ☐ vähintään 20 tuntia viikossa (kokopäivätyö)
- ☐ alle 20 tuntia viikossa (osapäivätyö)
- ☐ vaihtelevasti
- ☐ olen koululainen/opiskelija
- ☐ olen eläkkeellä
- ☐ olen tällä hetkellä työtön
- ☐ olen varusmies/siviilipalvelusmies
- ☐ en ole ansiotyössä tällä hetkellä (esim. kotiäiti tai -isä)

4. Ammatti (jos käytte töissä)

5. Ajokortti

☐ On ☐ Ei

6. Onko teillä pääkaupunkiseudun matkakortti?

- ☐ Ei ole matkakorttia
- ☐ On, valitkaa mitä sille on ladattu:
- ☐ Ladattu vain arvoa (rahaa)
- ☐ Ladattu sisäistä kautta, minkä kunnan? _____
- ☐ Ladattu usean kunnan kattavaa kautta (seutu, lähiseutu)
- ☐ Muu (esim. vapaalippu)

7. Kuinka usein teillä on polkupyörä käytettävissänne?

- ☐ Aina tai melkein aina
- ☐ Silloin tällöin
- ☐ Hyvin harvoin
- ☐ Ei ollenkaan

8. Mitä kulkutapaa (tai kulkutapojen yhdistelmää) useimmin käytätte työ-/koulumatkalla?

(Jos käytätte samalla matkalla useampaa kulkuvälinettä, esimerkiksi menette autolla junalle tai bussilla metrolle, kirjatkaa kaikki käytetyt kulkuvälineet.)

Jatkuu kääntöpuolella →

9. Kuinka usein käytätte seuraavia kulkutapoja:

	Päivittäin	Useita kertoja viikossa	Muutaman kerran kuussa	Harvemmin	En koskaan
Kävely (koko matka)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Polkupyörä	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Auto	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bussi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Raitiovaunu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Juna	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Metro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10. Kotiosoitteenne

Lähiosoite

Kaupunginosa

Kunta

11. Työ-/opiskelupaikan osoite

Lähiosoite

Kaupunginosa

Kunta

12. Jos viette säännöllisesti lapsia kouluun tms. niin lapsen koulun/päivähoitopaikan osoite

Lähiosoite

Kaupunginosa

Kunta

13. Useimmin ruokaostokseen käyttämänne kauppa

Kauppan nimi (tai kauppakeskus)

Lähiosoite

Kaupunginosa

Kunta

14. Toiseksi useimmin ruokaostokseen käyttämänne kauppa

Kauppan nimi (tai kauppakeskus)

Lähiosoite

Kaupunginosa

Kunta

15. Onko muita paikkoja, joissa käytte useita kertoja viikossa?
(1) Käynnin tarkoitus

Paikka

Lähiosoite

Kaupunginosa

Kunta

(2) Käynnin tarkoitus

Paikka

Lähiosoite

Kaupunginosa

Kunta

Matkapäiväkirja

Matka on siirtyminen paikasta A paikkaan B, esimerkiksi kotoa kauppaan. Jos kuljet päivän aikana kotoa töihin, töistä kauppaan ja kaupasta kotiin, olet tehnyt kolme matkaa. Myös ulkoilu tai esimerkiksi lenkki koiran kanssa on matka, vaikka se alkaisi ja kotoa ja päättyisi kotiin.

Tässä tutkimuksessa halutaan saada selvästi esiin eri kulkutavoilla tehdyt siirtymät, joten yksi matka voi koostua useammasta osuudesta, jos käytät sen aikana useampia kulkutapoja. (esim. autolla kauppaan = 1 matka, joka koostuu vain yhdestä osuudesta, mutta kävellen bussipysäkillä, sitten bussilla toiselle pysäkillä, josta kävellen töihin = 1 matka, joka koostuu kolmesta osuudesta.)

Lomakkeeseen täytetään yksi rivi jokaista matkan osuutta kohti, eli **Yksi rivi aina kun vaihdatte kulkumuotoa**. Täyttäkää uusi lomake jokaista tutkimuspäivää kohti.

Lähtö- ja saapumisajat ja saapumispaikka sekä **käytetty kulkumuoto** täytetään jokaista matkan osuutta kohti (ks. esimerkki). **Matkan viimeinen saapumispaikka**, eli se johon koko ajan on oltu matkalla, kirjataan mahdollisimman tarkasti (**osoite** ja kaupunki jos tiedossa (esimerkiksi **kuitista** löytyy usein kaupan, ravintolan tms. osoite), muuten mahdollisimman yksilöivä tunnistus jonka avulla paikan voi tunnistaa myöhemmin, esim. rakennuksen tai liikkeen nimi ja kaupunginosa, kuten Tennispalatsi, Malmitalo, Tapiolan Stockmann, Leppävaaran urheilupuisto). Kohtaan **saapumispaikan tyyppi** kirjataan vielä matkan lopullisen kohteen tyyppi takakannessa olevan listan A mukaisesti. Kulkutavanvaihtopaikat voi kirjata vähemmän tarkasti, esim. bussipysäkki Kannelmäessä. **Matkan tarkoituksen ja säätilan** voi täyttää kerran matkaa kohti, viimeisen kulkutavan kohdalle. Matkan tarkoitus valitaan vaihtoehdoista, jotka on listattu takakannessa (lista B). **Säätilan** kohdalle merkitään myös jos keli on huono, esim. liukas tai sohjoinen.

Jos matka sisältää vaikkapa bussin tai junan odottelua pysäkillä, sitä ei tarvitse kirjata erikseen, jos se käy helposti ilmi muutenkin, esimerkiksi esimerkissä klo 8.29-8.31 ja 8.45-8.46. Jos kuitenkin joudutte vaihtaessanne kulkuvälineestä toiseen kävelemään pysäkillä toiselle, kirjatkaa aikaväli lomakkeeseen ja matkan tarkoituksen kohdalle **vaihto**, koska GPS-laite rekisteröi tällöin liikettä (ks. esimerkki klo 8.43-8.45).

Kohtaan **muita huomioita ja ongelmia** voitte kirjata muita matkan sujumiseen vaikuttavia tekijöitä (esimerkiksi ruuhka), tai jos teillä on mukanaan esimerkiksi lastenvaunut, paljon kantamuksia tai muuta, mikä vaikuttaa matkantekoon.

Kääntöpuolella on **esimerkki lomakkeen täyttämisestä**. Esimerkissä on korostettu **keltaisella** matkojen lähtö- ja saapumispaikat, joissa pitää painaa GPS-laitteen nappulaa (katsokaa lisätietoa GPS-ohjeesta). Lisäksi ensimmäiseen sarakkeeseen merkitään, milloin nappulaa on muistettu painaa.

Takakannessa on listat matkan tarkoituksenvaihtoehdoista ja saapumispaikan tyypeistä.

Saapumispaikoista voitte käyttää esimerkiksi seuraavia **lyhenteitä**. Jos käytätte muita lyhenteitä, muistakaa selittää ne.

BP = bussipysäkki (kirjatkaa lisäksi bussipysäkin nimi, numero tai sijainti muuten)

M = metroasema (kirjatkaa lisäksi metroaseman nimi)

Esimerkki kääntöpuolella →

1

ESIMERKKI

Pvm. 5.1.2009 Viikonpäivä maanantai Päivän ensimmäisen matkan lähtöosoite/-paikka Koti

Painoin nappulaa	Lähtö-aika	Käytetty kulkumuoto + bussin nro / junan kirjain tai vastaava	Saapumisosoite / -paikka (ja kaupunginosa / kaupunki)	Saapumispaikan tyyppi (valitkaa takasivun listasta A)	Saapumis-aika	Matkan tarkoitus (valitkaa takasivun listasta B)	Arvio säästä matkan aikana (n. °C, pilvisyys, sade, ajo-/kävelykeli)	Muita huomioita tai ongelmia
x	8.19	Kävely	Bussipysäkki, Leppävaara	-	8.29	-	-	Esimerkki: nämä viisi osuutta muodostavat yhden matkan. Bussipysäkkistä tarvitaan nimi, numero tai kaupunginosa. Lyhenteitä voi käyttää, esim. BP=bussipysäkki, M=metro
	8.31	Bussi (linja 550)	Bussipysäkki, Tapiola	-	8.43	-	-	
	8.43	Kävely	Bussipysäkki Tapiola	-	8.45	Vaihto	-	
	8.46	Bussi (linja 501)	BP (Kivaidankatu), Lauttasaari	-	8.55	-	-	
x	8.55	Kävely	Työpaikka	3 (oma työpaikka)	9.01	1 (työmatka)	-5 °C, aurinkoista, liukasta	
Jättäkää tyhjä rivi eri matkojen väliin								
x	16.02	Auto	Simonkatu 5, Helsinki	-	16.22	-	-	Ruuhkaa
	16.22	Kävely	Helsingin rautatieasema	-	16.29	Vaihto	-	
	16.32	Juna, A	Pitäjänmäen asema	-	16.44	-	-	
x	16.44	Kävely	Pitäjänmäentie 12, Helsinki	16 (vierailupaikka)	16.46	11 (vierailu)	-3 °C, puoli-pilvistä, sohjoa	
Jättäkää tyhjä rivi eri matkojen väliin								
x	18.15	Kävely	Kauppakeskus Sello, Leppävaara	10 (kauppakeskus)	18.37	4 (päivittäistavaroiden osto)	-7 °C, pilvistä, sohjoa	
Jättäkää tyhjä rivi eri matkojen väliin								
x	19.05	Kävely	Koti	1 (oma koti)	19.12	4	-7 °C, pilvistä	Paljon kauppakeskeä
x								

2

Pvm. _____ Viikonpäivä _____ Päivän ensimmäisen matkan lähtöosoite/-paikka _____

Painoin napakkeita	Lähtö- aika	Käytetty kulkumuoto + bussin nro / junan kirjain tai vastaava	Saapumisosoite / -paikka (ja kaupunginosa / kaupunki)	Saapu- mispaikan tyyppi (valitkaa takasivun listasta A)	Saapu- mis- aika	Matkan tarkoitus (valitkaa takasivun listasta B)	Arvio säästä matkan aikana (n. °C, pilvisyys, sade, ajo-/kävelykeli)	Muita huomioita tai ongelmia

Painoin nappulaa	Lähtö- aika	Käytetty kulkumuoto + bussin nro / junan kirjain tai vastaava	Saapumisoite / -paikka (ja kaupunginosa / kaupunki)	Saapu- mispaikan tyyppi (valitkaa takasivun listasta A)	Saapu- mis- aika	Matkan tarkoitus (valitkaa takasivun listasta B)	Arvio säästä matkan aikana (n. °C, pilvisyys, sade, ajo-/kävelykeli)	Muita huomioita tai ongelmia

GPS-tutkimus – Loppukyselylomake

Tämän viimeisen kyselyn tarkoituksena on kartoittaa, millainen kuva osallistujille jäi tutkimuksesta ja siihen osallistumisesta. Erityisesti olisi hyvä tietää, miten rasittavana tutkimukseen osallistuminen koettiin. Sekä risut että ruusut auttavat tutkimuksen onnistumisen arvioinnissa ja joskus mahdollisesti toteutettavan laajemman tutkimuksen suunnittelussa.

Jos ette halua vastata joihinkin kysymyksiin, voitte jättää ne tyhjiksi. Toivon kuitenkin vastausta mahdollisimman moneen kysymykseen, koska osallistujien joukko on niin pleni.

Kiitos vastauksista, ja erityisen paljon kiitoksia osallistumisesta tutkimukseen.

A) Tausta vastaajasta ja tutkimukseen osallistumisesta

1. Ikä _____ vuotta

2. Sukupuoli

☐ Nainen ☐ Mies

3. Millä kulkutavoilla teitte suurimman osan matkoistanne tutkimuspäivinä? (voitte valita useampia)

a) Päivä 1:

☐ Kävelen

☐ Polkupyörällä

☐ Autolla

☐ Bussilla

☐ Raitiovaunulla

☐ Junalla

☐ Metrolla

☐ Muulla, millä?

b) Päivä 2:

☐ Kävelen

☐ Polkupyörällä

☐ Autolla

☐ Bussilla

☐ Raitiovaunulla

☐ Junalla

☐ Metrolla

☐ Muulla, millä?

4. Oliko teillä auto käytettävissänne tutkimuspäivinä?

a) Päivänä 1: ☐ Kyllä ☐ Ei

b) Päivänä 2: ☐ Kyllä ☐ Ei

5. Mitkä olivat työ-/opiskeluaikanne tutkimuspäivinä?

a) Päivä 1: Alkoi klo _____
Päättyi klo _____

b) Päivä 2: Alkoi klo _____
Päättyi klo _____

B) Tutkimuksen rasittavuus ja ongelmat

6. Mitä mieltä olette tutkimuksen kestosta?

☐ Liian pitkä, osallistuminen oli rasittavaa

☐ Sopiva, en olisi halunnut jatkaa enää

☐ Olisin voinut jatkaa vielä

Muita kommentteja _____

7. a) Kuinka pitkään samanlaiseen tutkimukseen voisitte kuvitella osallistuvanne?

☐ 1-2 päivää

☐ 3-5 päivää

☐ Viikko

☐ Kaksi viikkoa

☐ Kuukausi

☐ En haluaisi osallistua

Jatkuu kääntöpuolella →

7. b) Kuinka pitkään vastaavaan tutkimukseen voisitte kuvitella osallistuvanne, jos teidän ei tarvitsisi täyttää käsin paperista matkapäiväkirjaa, vaan ainoastaan kuljettaa GPS-laitetta mukanaan ja muistaa ladata se joka ilta.

- ☐ 1-2 päivää
☐ 3-5 päivää
☐ Viikko
☐ Kaksi viikkoa
☐ Kuukausi
☐ En haluaisi osallistua

Jos vastasitte a) tai b)-kohtaan tai molempiin, että ette haluaisi osallistua, miksi ette?

8. Mikä oli rasittavinta tutkimuksessa? (voitte valita useita kohtia)

- ☐ Matkapäiväkirjan täyttäminen
☐ GPS-laitteen kuljettaminen mukana
☐ GPS-laitteen lataaminen
☐ En kokenut tutkimusta rasittavaksi

Selittäkää tässä tarkemmin, mistä rasitus ja vaiva aiheutuivat:

9. Oliko teillä tutkimuksen aikana vaikeuksia jossain seuraavista (voitte valita useita):

- ☐ Matkapäiväkirjan täyttäminen
☐ GPS-laitteen lataaminen
☐ GPS-laitteen käyttäminen
☐ GPS-laitteen kuljettaminen mukana
☐ GPS-laitteen muistaminen ulos lähtiessä
☐ GPS-signaalin löytyminen (jos kiinnitit siihen huomiota)
☐ Muita ongelmia

Selittäkää tässä tarkemmin tai kertokaa muista ongelmista:

- C) Vapaata palautetta tutkimuksesta

Liite 4. Kenttätutkimuksen osallistujille jaettu ohjeistus (3 sivua) tutkimuksen suorittamisesta ja GPS-laitteen käyttämisestä.

Pikaohje GPS-paikannuslaitteen käyttöön

Muistakaa myös täyttää **matkapäiväkirjaa**! Tarkoituksena on, että paikannuslaitteen muistista ja täyttämästänne matkapäiväkirjasta löytyy samat matkat.

Pitäkää GPS-laitetta lähellä avaimia, kännykkää tms. jotta muistatte pitää sitä mukana.

- **Aamulla** lähtiessänne ulos laittakaa laite päälle eli liukukytkin asentoon **LOG** (oranssi valo syttyy ja alkaa hetken kuluttua vilkkua).
- **Painakaa nappulaa**, joka on laitteessa valojen keskellä aina, kun aloitatte uuden matkan tai saavutte määränpäähän (laite tallentaa "erikoismerkinnän"). Jos unohdatte painaa nappulaa, jatkakaa kuitenkin tutkimusta normaalisti.
- Pitäkää laitetta mukana ja päällä koko päivän ajan. Rakennusten sisällä laite ei yleensä saa signaalia (jolloin oranssi valo palaa koko ajan), mutta tämä ei haittaa. Tultuanne takaisin ulos valon pitäisi taas hetken kuluttua alkaa vilkkua.
- (Jos menette **sairaalaan/terveyskeskukseen tai lentokoneeseen**, laittakaa laite pois päältä laittamalla kytkin asentoon **OFF**. Tullessanne takaisin ulos laittakaa kytkin taas asentoon **LOG**.)
- **Illalla** kun tulette kotiin, **ettekä ole enää lähdössä mihinkään**, laittakaa kytkin asentoon **OFF** ja laite lataukseen (vihreä valo syttyy). Pitäkää laite latauksenkin aikana mahdollisimman näkyvässä paikassa, jotta muistatte ottaa sen mukaan taas seuraavana aamuna.

GPS-laitteen akku ei kestä yli kymmentä pakkasastetta, eli **jos pakkasta on yli 10 astetta, älkää käyttäkö laitetta** (tällöin matkapäiväkirjaakaan ei tarvitse täyttää). Jos pakkas laskee päivän aikana, laittakaa laite päällä ja ruvetkaa täyttämään myös matkapäiväkirjaa. Jos kova pakkas jatkuu koko päivän, siirretään tutkimus toiselle päivälle.

Laitetta voi pitää taskussa tai laukussa, mutta varmistakaa, ettei laitteen ja kankaan välissä ole muita esineitä ja ettei tasku ole niin tiukka, että laite joutuu puristuksiin. Jos lähdette liikkeelle ilman taskuja tai laukkuja, esim. lenkille, ja laitteen ottaminen mukaan on liian vaikeaa, voitte jättää sen kotiin. Jos ette halua jonkin matkan tallentuvan, laittakaa kytkin kyseisen matkan ajaksi asentoon **OFF**, tai älkää ottako laitetta mukaan.

Jos tulee mitään kysyttävää tai ongelmia, soittakaa Annikalle numeroon 040-8351357 (mihin vuorokauden aikaan tahansa). Jos jostain syystä en vastaa, älkää huolestuko vaan yrittäkää myöhemmin uudelleen.

Jos vaikuttaa siltä, että GPS-laite jostain syystä ei toimi kunnolla (esimerkiksi jos laite ei löydä satelliitteja, eli oranssi valo syttyy, mutta ei moneen minuuttiin rupea vilkkumaan)

- varmistakaa, että laitteella on "**näköyhteys**" taivaalle eli laitteen ja taivaan välissä ei ole esteitä, ja odottakaa muutama minuutti (nopeasti liikkuvassa välineessä satelliittien löytyminen voi kestää useita minuutteja). Jos tämä ei auta
- laittakaa kytkin asentoon **OFF** ja muutaman sekunnin kuluttua takaisin asentoon **LOG**. Odottakaa muutama minuutti. Jos tämä ei auta
- laittakaa kytkin asentoon **OFF** ja antakaa olla siinä. Vasta kun aloitatte seuraavan matkan, laittakaa kytkin asentoon **LOG**. Jos tämä ei auta
- ladatkaa laitetta jonkin aikaa ja yrittäkää uudelleen. Jos tämäkään ei auta
- soittakaa Annikalle yllä mainittuun numeroon.

Sisätiloissa laite yleensä ei löydä satelliitteja, ja oranssi valo palaa koko ajan. Tämä on täysin normaalia eikä haittaa tutkimusta! Samoin esimerkiksi liikennevälineissä tai korkeiden rakennusten välissä signaali voi hävitä ja oranssi valo palaa koko ajan. Tästä ei pidä huolestua, satelliitit löytyvät taas kun palaatte parempiin olosuhteisiin.

Tarkennuksia pikaohjeeseen:

Pitäkää laite päällä (LOG-asennossa) koko ajan aamusta iltaan sekä tekemillänne matkoilla että niiden välissä sisätiloissa (ei tarvitse yöllä kun olette kotona)

- **paitsi sairaalassa/terveyskeskuksessa ja lentokoneessa**

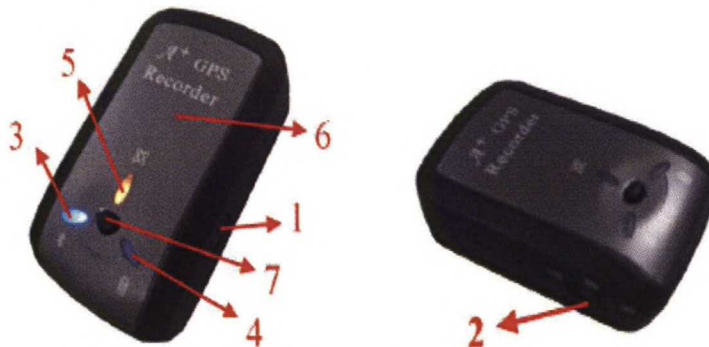
Eli (aamulla) kun lähdette kotoa, laittakaa sivussa oleva kytkin kohtaan **LOG** (oranssi valo syttyy ja alkaa jonkin ajan kuluttua vilkkua). Aina kun aloitatte uuden matkan tai saavutte määränpäähän, painakaa laitteessa valojen keskellä olevaa nappulaa (oranssi satelliittien merkkivalo vilkkuu tällöin hetken punaisena). Nappulan painaminen vastaa siis matkapäiväkirjaan merkittyjä matkoja (jotka koostuvat useista osista). Kun vaihdatte kulkutapaa matkan aikana, kirjatkaa uusi rivi matkapäiväkirjaan, mutta älkää painako laitteen nappulaa.

Illalla kotiin tultuanne, kun ette enää ole lähdössä mihinkään, laittakaa kytkin asentoon **OFF** ja **laite lataukseen**. (Kun lataus aloitetaan, syttyy vihreä valo. Kun akku on täynnä, vihreä valo sammuu, ja latauksen voi lopettaa.)

Kun kytkimen laittaa LOG-asentoon, laitteeseen pitäisi ensin syttyä kiinteä oranssi valo. Kun valo alkaa vilkkua, laite on löytänyt satelliitit eli se alkaa tallentaa reittiä. Tämä voi kestää hetken, älkää huolestuko. **Normaalissa tallennustilassa (ollessanne ulkona, liikennevälineessä tms.) laitteessa vilkkuu oranssi valo (☼), ja muut valot eivät pala!** (Jos muita valoja palaa, katsokaa seuraavalta sivulta mikä on ongelma.)

(HUOM! Jos kytkin on NAV-asennossa, laite paikantaa sijainnin, mutta ei tallenna sitä. Tätä asentoa ei pidä käyttää tässä tutkimuksessa. Jos vahingossa laitatte kytkimen NAV-asentoon, laitteeseen syttyy sininen valo. Jos näin tapahtuu, laittakaa kytkin asentoon OFF ja sitten takaisin asentoon LOG.)

1.5 Appearance




1. Lataus tai tiedonsiirto (mini USB liitäntä)
2. Toiminnanvalitsin: **OFF**(virta pois) / **NAV**(navigointi) / **LOG**(navigointi ja sijainnin tallennus)
3. Bluetoothin tila, LED-valo (sininen)
4. Akun tila, LED-valo (punainen/vihreä)
5. GPS:n tila, LED-valo (oranssi) / Paikan tallennus, LED-valo (punainen)
6. Sisäänrakennettu GPS-antenni
7. Painonappi, tallentaa senhetkisen sijainnin (valo 5 vilkkuu punaisena hetken)


Valojen merkitys:




Normaalikäytössä reittiä tallennettaessa LOG-tilassa pitäisi ensin syttyä kiinteä oranssi valo, ja kun laite on löytänyt satelliitit, oranssi valo alkaa vilkkua. Jos muita valoja palaa, se on merkki ongelmasta, paitsi jos painonappia on painettu tarkoituksella tai vahingossa (tällöin punainen valo vilkkuu kolme kertaa).

Punainen valo  vilkkuu : akku vähissä -> lataa akku heti kun mahdollista.

Punainen valo  vilkkuu hitaasti jatkuvasti: muisti alkaa täyttyä (tälle ette voi tehdä mitään -> ei tarvitse huolestua jos on jo toisen tutkimuspäivän ilta, mutta jos valo alkaa vilkkua aikaisemmin).

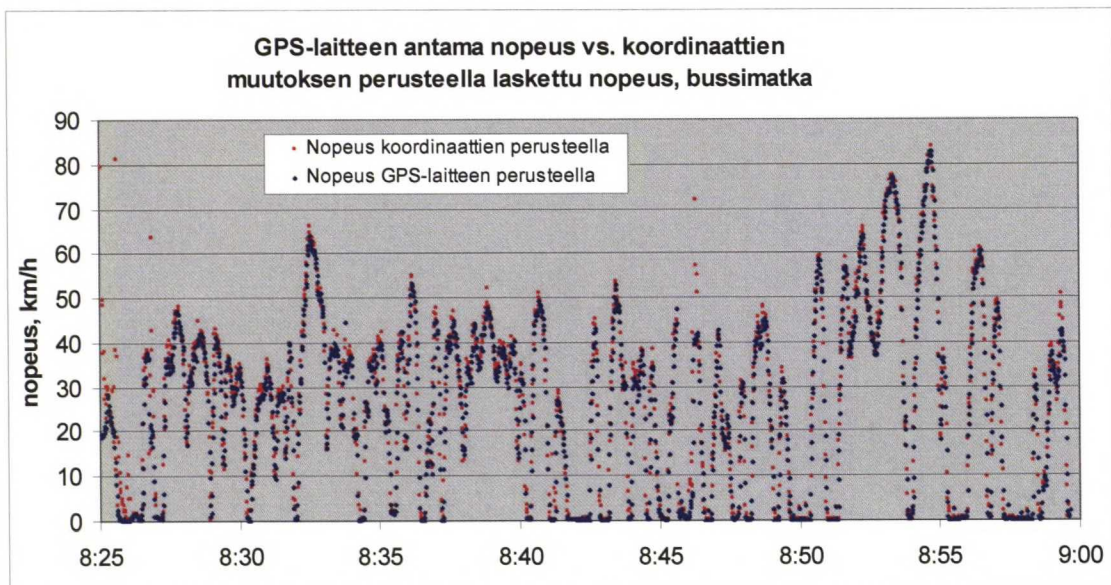
Sininen valo  : laite on NAV-tilassa (ei tallenna sijaintia) -> laita kytkin asentoon LOG

Vihreä valo  : Kun lataus aloitetaan, syttyy vihreä valo. Kun akku on täynnä, vihreä valo sammuu, ja latauksen voi lopettaa.

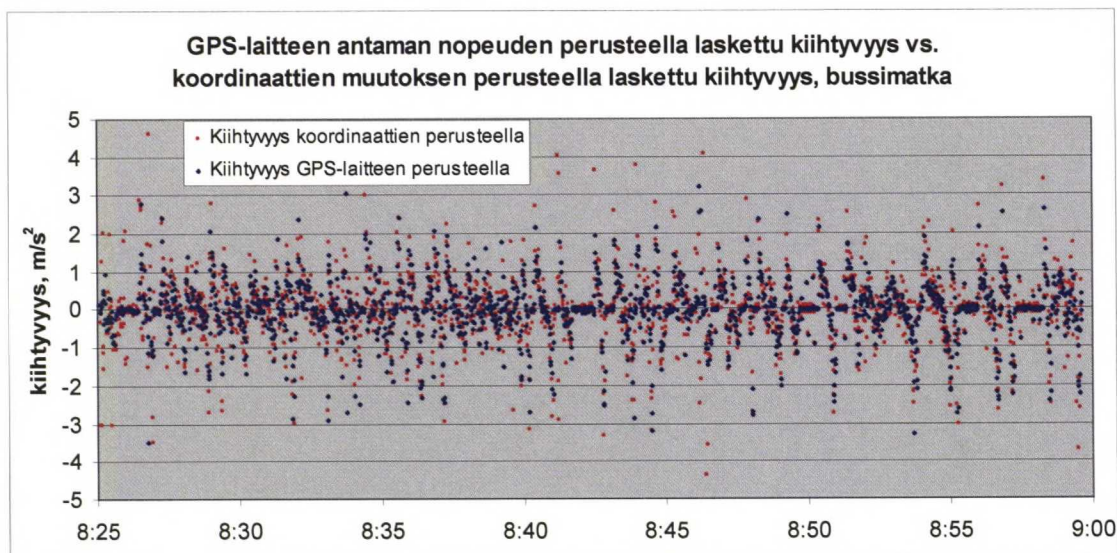
Kategoria	Symboli	Väri	Valon tila	Selitys
Bluetooth-yhteyden tila, LED-valo		Sininen ●	Palaa koko ajan	Bluetooth on päällä, mutta laitteella ei ole yhteyttä toiseen Bluetooth-laitteeseen (NAV-tila)
			Vilkkuu hitaasti (1 kerran 5:ssä sekunnissa)	Nukkumistila
			Vilkkuu nopeasti (1 kerran 2 sekunnissa)	Bluetooth yhteys toiseen laitteeseen on muodostettu ja laite on valmis datan siirtoon (NAV-tila)
Akun tila, LED-valo		Punainen ●	Vilkkuu	Akku on lähes tyhjä.
		Vihreä ●	Palaa koko ajan	Akun lataus on käynnissä (valo syttyy kun latauksen aloittaa)
		Vihreä ●	Ei pala	Akku on täyteen ladattu
GPS-paikannuksen tila, LED-valo		Oranssi ●	Palaa koko ajan	Laite etsii satelliitteja, sijainti ei ole vielä selvillä (valo syttyy heti kun laitteen laittaa NAV- tai LOG-tilaan)
			Vilkkuu nopeasti	Satelliitit on löydetty ja laite kertoo sijainnin
		Punainen ●	Vilkkuu nopeasti	LED vilkkuu kolme kertaa, kun painonappia painetaan kiinnostavan sijainnin tallentamiseksi
			Vilkkuu hitaasti	Muistia on tyhjänä alle 20%.
			Palaa koko ajan	Muisti on täynnä (sijaintia ei enää tallenneta (tai vanhoja pyyhkiytyy pois)).

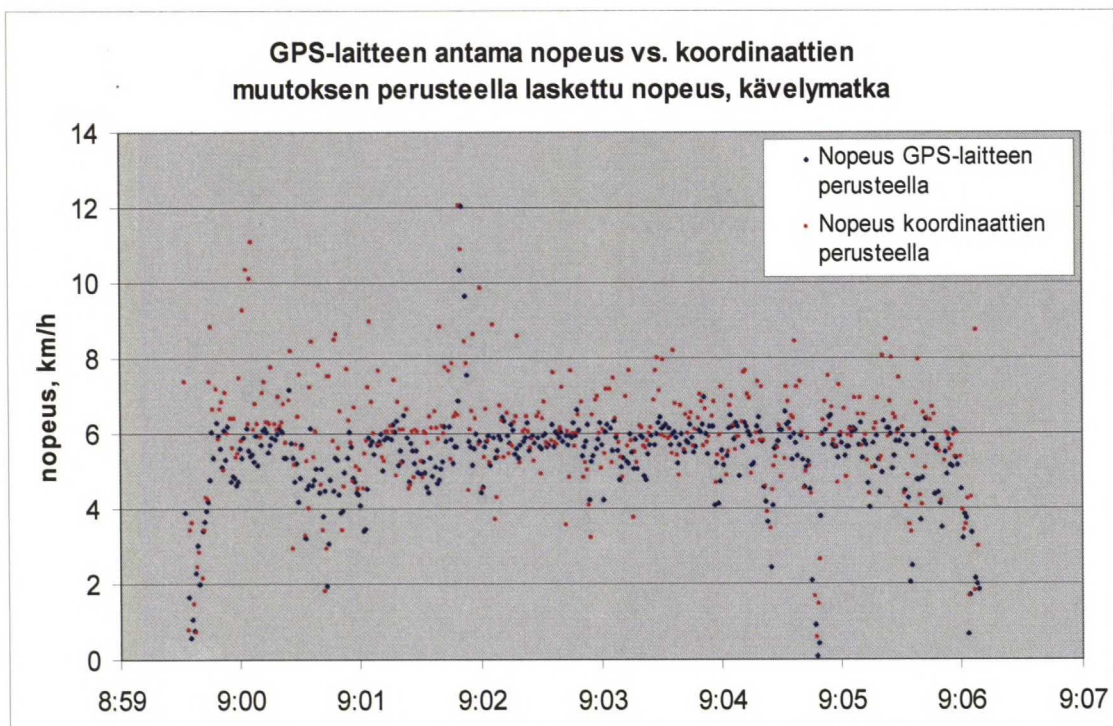
Liite 5. Nopeus- ja kiihtyvyysskuvaajia, joissa on esitetty sekä suoraan GPS-laitteesta saatu nopeus että koordinaattien muutoksen perusteella laskettu nopeus, tai vastaavasti näiden pohjalta lasketut kiihtyvyydet. (Koordinaattien perusteella laskettu nopeus punaisella, GPS-laitteesta saatu nopeus sinisellä.)

Kuvaajista näkyy, miten rikkonaista matkanteko on bussilla tai autolla kaupunki-alueella; nopeus ei pysy vakiona pitkää aikaa, vaan jatkuvasti pitää kiihdyttää ja jarruttaa, usein jopa pysähtyä. Ajaminen moottoritillä tai vastaavalla etuajo-oikeutetulla tiellä, jossa nopeusrajoitus on korkea, erottuu selvästi automatkan kuvaajasta.

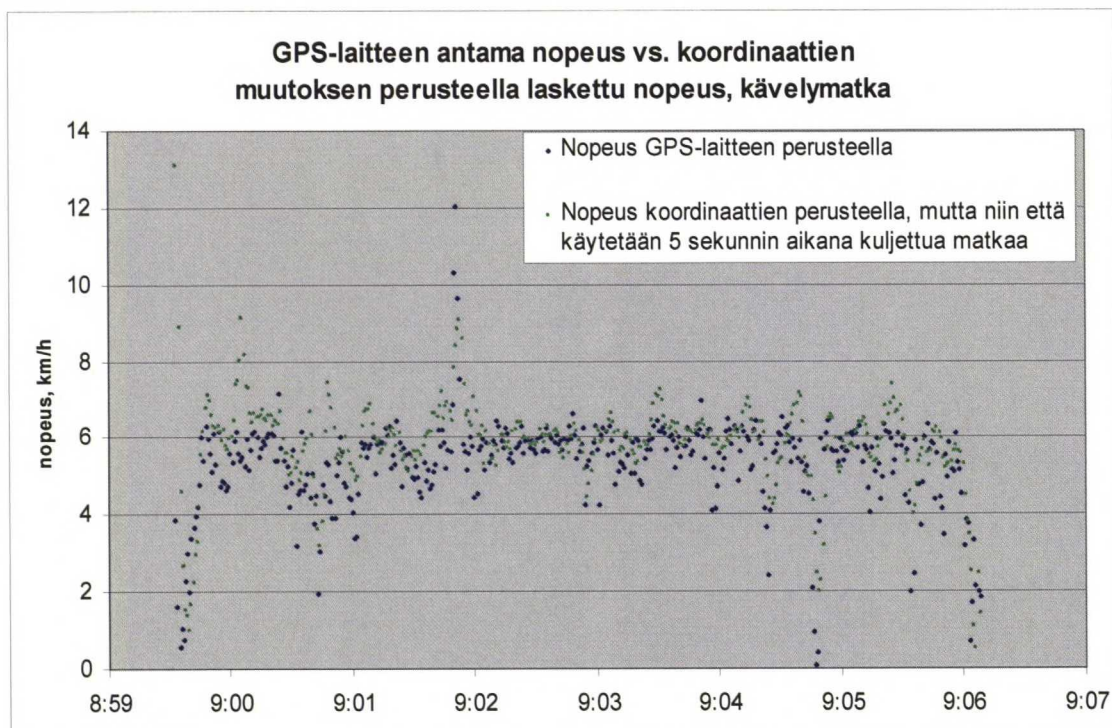


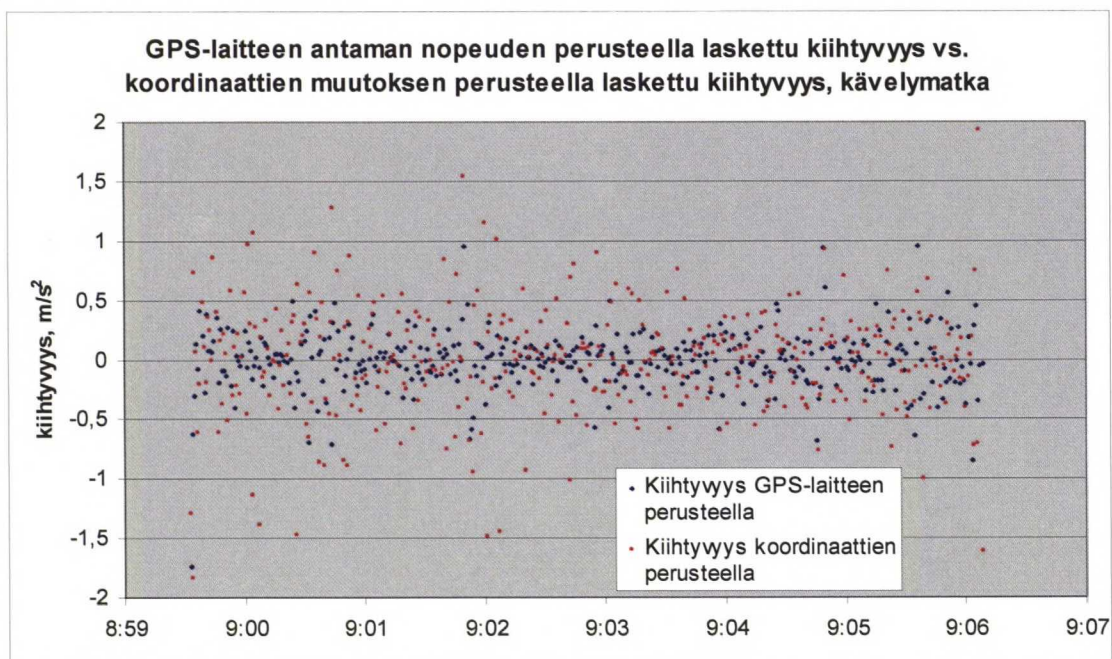
Poikkeukselliset pisteet koordinaattien perusteella laskettujen nopeuksien joukossa johtuvat luultavasti paikannusvirheistä. Kun peräkkäisten pisteiden välinen etäisyys on väärä, myös hetkellinen nopeus saa poikkeavan arvon.

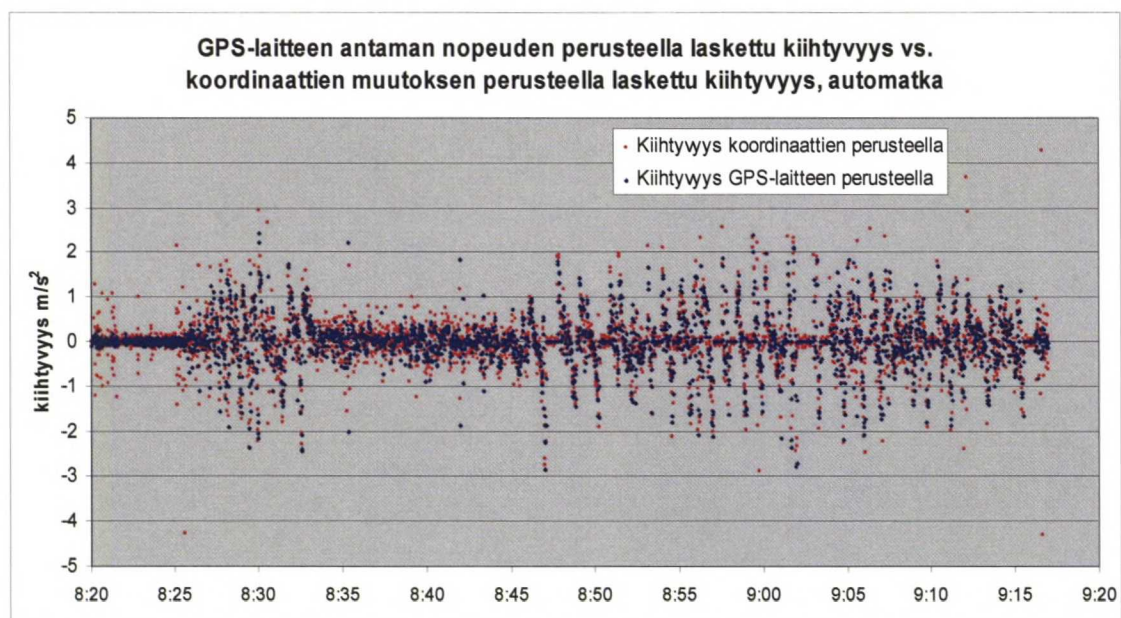
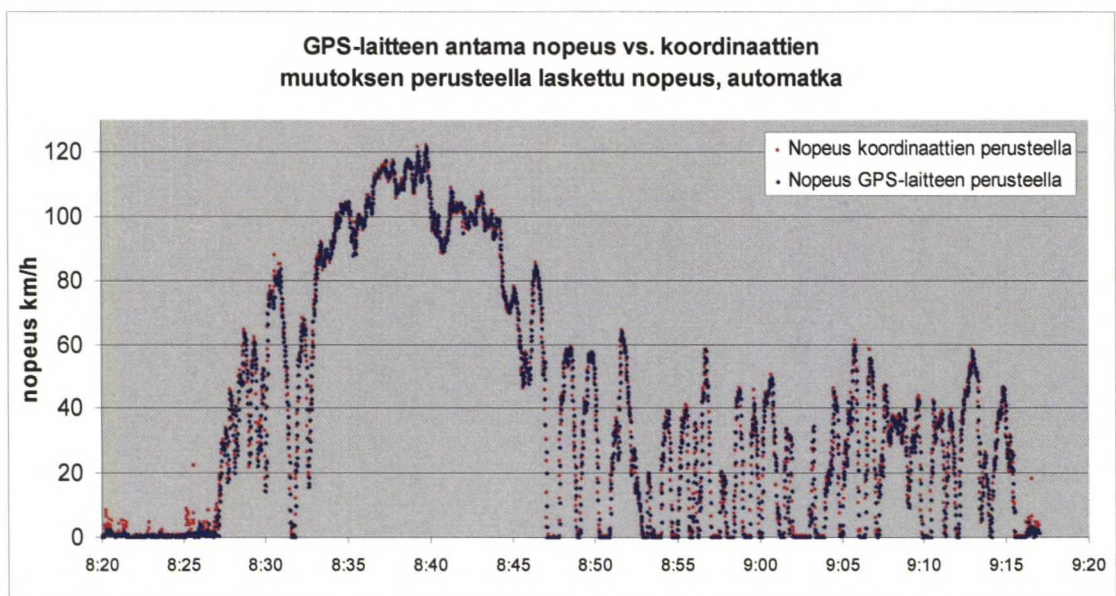




Kävelyssä koordinaattien perusteella laskettu nopeus näyttää korkeammalta ja vaihtelee enemmän kuin GPS-laitteen antama. Tämä saattaa johtua siitä, että koordinaatit on tallennettu sekunnin välein, ja kävelijä liikkuu niin lyhyen matkan yhden sekunnin aikana, että siitä laskettu nopeustieto ei ole luotettava. Alla on sama kuva kuin yllä, mutta nyt koordinaattien perusteella lasketun nopeuden määrittämisessä on käytetty edellisten 5 sekunnin aikana kuljettua matkaa yhden sekunnin aikana kuljetun matkan sijaan. Tulokset ovat näin jo paljon yhtenevämpiä kuin yllä.







Liite 6. Versioita päättelysääntöihin perustuvasta kulkutavanpäättelyalgoritmista

Kulkutavanpäättelyalgoritmin versio 0.0. Mahdollisimman yksinkertainen, keskinopeuksiin perustuva algoritmi.

1) Jos nopeuksien keskiarvo on alle 8 km/h, kulkutapa on kävely.
2) Jos nopeuksien keskiarvo on 8–20 km/h, kulkutapa on pyöräily.
3) Jos nopeuksien keskiarvo on 20–25 km/h, kulkutapa on raitiovaunu.
4) Jos nopeuksien keskiarvo on 25–45 km/h, kulkutapa on bussi.
5) Jos nopeuksien keskiarvo on 45–60 km/h, kulkutapa on metro.
6) Jos nopeuksien keskiarvo on 60–80 km/h, kulkutapa on auto.
7) Jos nopeuksien keskiarvo on 80–160 km/h, kulkutapa on juna.

Kulkutavanpäättelyalgoritmin versio 1.0. Lisätietona joukkoliikenteen pysäkkien paikat

1) Jos nopeuksien keskiarvo on alle 8 km/h, kulkutapa on kävely.
2) Jos nopeuksien keskiarvo on 8–25 km/h, niin <ul style="list-style-type: none">- jos alku- tai loppupiste osuu raitiovaunupysäkille, niin kulkutapa on raitiovaunu- jos alku- tai loppupiste osuu bussipysäkille, niin kulkutapa on bussi- muutoin kulkutapa on pyöräily
3) Jos nopeuksien keskiarvo on 25–40 km/h, niin <ul style="list-style-type: none">- jos alku- tai loppupiste osuu bussipysäkille, niin kulkutapa on bussi- muutoin kulkutapa on auto
4) Jos nopeuksien keskiarvo on 40–70 km/h, niin <ul style="list-style-type: none">- jos alku- tai loppupiste osuu metroasemalle, niin kulkutapa on metro- jos alku- tai loppupiste osuu bussipysäkille, niin kulkutapa on bussi- muutoin kulkutapa on auto
5) Jos nopeuksien keskiarvo on 70–120 km/h, kulkutapa on auto.
6) Jos nopeuksien keskiarvo on yli 120 km/h, kulkutapa on juna.
7) Etsitään vielä erikseen juna, jotka ovat jääneet huomiotta: <ul style="list-style-type: none">- jos nopeuksien keskiarvo on yli 30 km/h, ja alku- tai loppupiste osuu juna-asemalle, niin kulkutapa on juna

Kulkutavanpäättelyalgoritmin versio 1.11. Lisätietona joukkoliikenteen pysäkkien paikat, nopeuden 90. persentiili ja nopeuden keskihajonta.

1) Jos nopeuksien keskiarvo on $8 < \text{km/h}$, kulkutapa on kävely.
2) Jos nopeuksien keskiarvo on 8–22 km/h, niin <ul style="list-style-type: none">- jos alku- tai loppupiste osuu raitiovaunupysäkille ja nopeuden 90. persentiili $> 15 \text{ km/h}$ ja nopeuden keskihajonta 6–20 km/h, niin kulkutapa on raitiovaunu- jos alku- tai loppupiste osuu bussipysäkille ja nopeuden 90. persentiili $> 30 \text{ km/h}$ ja nopeuden keskihajonta $> 10 \text{ km/h}$, niin kulkutapa on bussi- jos nopeuden 90. persentiili $< 25 \text{ km/h}$ ja nopeuden keskihajonta $< 10 \text{ km/h}$, niin kulkutapa on pyöräily- muutoin kulkutapa on auto
3) Jos nopeuksien keskiarvo on 22–40 km/h, niin <ul style="list-style-type: none">- jos alku- tai loppupiste osuu bussipysäkille, niin kulkutapa on bussi- muutoin kulkutapa on auto
4) Jos nopeuksien keskiarvo on 40–70 km/h, niin <ul style="list-style-type: none">- jos alku- tai loppupiste osuu metroasemalle, niin kulkutapa on metro- jos alku- tai loppupiste osuu bussipysäkille, niin kulkutapa on bussi- muutoin kulkutapa on auto
5) Jos nopeuksien keskiarvo on 70–120 km/h, kulkutapa on auto.

6) Jos nopeuksien keskiarvo on yli 120 km/h, kulkutapa on juna.
7) Etsitään vielä junia, jotka ovat jääneet huomiotta: <ul style="list-style-type: none"> - jos nopeuksien keskiarvo on yli 30 km/h, ja alku- tai loppupiste osuu juna-asemalle, niin kulkutapa on juna

Liite 7. Kulutavanpäättelyalgoritmin pisteytysjärjestelmä ja sen perustelut

Jos annettu kriteeri täyttyy, kulkutavalle annetaan taulukossa mainittu määrä pisteitä, ja kulkutavaksi valitaan se kulkutapa, jolla on eniten pisteitä. On myös hyvä olla säännöt, miten järjestys määräytyy, jos useammalla kulkutavalla on yhtä monta pistettä. Esimerkiksi tässä työssä on tehty todennäköisyyksien perusteella muutama sääntö siitä, mikä kulkutapa voittaa, jos kahdella kulkutavalla on yhtä paljon pisteitä, koska muuten asia ratkeaisi sen perusteella, millä numerolla mitäkin kulkutapaa ohjelmassa merkitään.

muuttuja	kriteeri	ratio-					
		kävely	pyörä	vaunu	bussi	metro	juna auto
keskinopeus	alle 6 km/h	3	1				
	6-8 km/h	1	1	1			
	8-25 km/h		1	1	1		1
	25-40 km/h				1		1
	40-70 km/h				1	1	1
	70-120 km/h						1
	yli 120 km/h					2	
nopeuden 90. persentiili	alle 30 km/h	1	1	1			
	30-60 km/h			1	1		1
	60-85 km/h				1	1	1
	85-100 km/h				1		1
	100-140 km/h					1	1
	yli 140 km/h					1	
nopeuksien keskihajonta	alle 6 km/h	1	1				
	6-10 km/h	1	1	1			
	10-20 km/h			1	1	1	1
	20-35 km/h				1	1	1
	yli 35 km/h					1	
kiihtyvyyden itseisarvojen keskiarvo	alle 0,3 m/s ²	1	1				1
	0,3-0,5 m/s ²			1	1	1	1
	yli 0,5 m/s ²				1	1	1
kiihtyvyyden 90. persentiili	alle 0,4 m/s ²	1	1				1
	0,4-0,6 m/s ²		1	1	1	1	1
	0,6-0,8 m/s ²			1	1	1	1
	0,8-1 m/s ²				1	1	1
	yli 1 m/s ²					1	
kiihtyvyyden keskihajonta	alle 0,4 m/s ²	1	1				
	0,4-0,6 m/s ²		1	1	1	1	1
	0,6-0,8 m/s ²				1	1	1
	yli 0,8 m/s ²					1	
lähtö- ja loppupisteet	lähellä ko. pysäkkiä ¹			1-2	1	1-2	1-2
	ei lähellä ko. pysäkkiä ²			-1	-2	-1	-1
matkalla tehdyt pysähdykset	yksikin pysähdys ko. pysäkillä ³			1	1	1	1
	useita pysähdyksiä ko. pysäkillä ⁴			1	1	1	1
	ei pysähdyksiä ko. pysäkillä ⁵			-1	-1	-1	-1

1) ensimmäinen tai viimeinen piste on lähellä ko. kulkutavan pysäkkiä; bussi saa joka tapauksessa yhden pisteen, ja muut joukkoliikennevälineet saavat yhden pisteen, jos matkan toinen pää on lähellä pysäkkiä ja kaksi pistettä, jos molemmat päät ovat lähellä pysäkkiä

2) jos kumpikaan matkan pää ei ole lähellä ko. kulkutavan pysäkkiä, vähennetään pisteitä

3) jos matkan aikana havaitaan yksikin pysähdys tai siltä näyttävä lähellä ko. kulkutavan pysäkkiä

4) jos pysähdyksiä lähellä pysäkkejä on useita, bussilla vähintään 10 ja muilla vähintään 5, niin annetaan lisäpiste edellisen kohdan pisteen lisäksi

5) jos pysähdyksiä lähellä ko. kulkutavan pysäkkejä ei havaita, vähennetään piste

Edellisen sivun pisteytystaulukko (samoin kuin päättelysäännötkin) on laadittu osin seuraavien kuvien pohjalta. Kuvissa on esitetty eri muuttujien vaihteluvälejä eri kulkutavoilla. Kuvat on tehty laskemalla kenttätutkimuksessa kerätyn GPS-aineiston jokaiselle matkan osalle kyseiset muuttujat ja analysoimalla muuttujien tunnuslukuja kulkutavoittain. Lasketut tunnusluvut olivat:

MAX = suurin arvo aineistossa

90 % = 90. persentiili (90 % arvoista on tätä pienempiä)

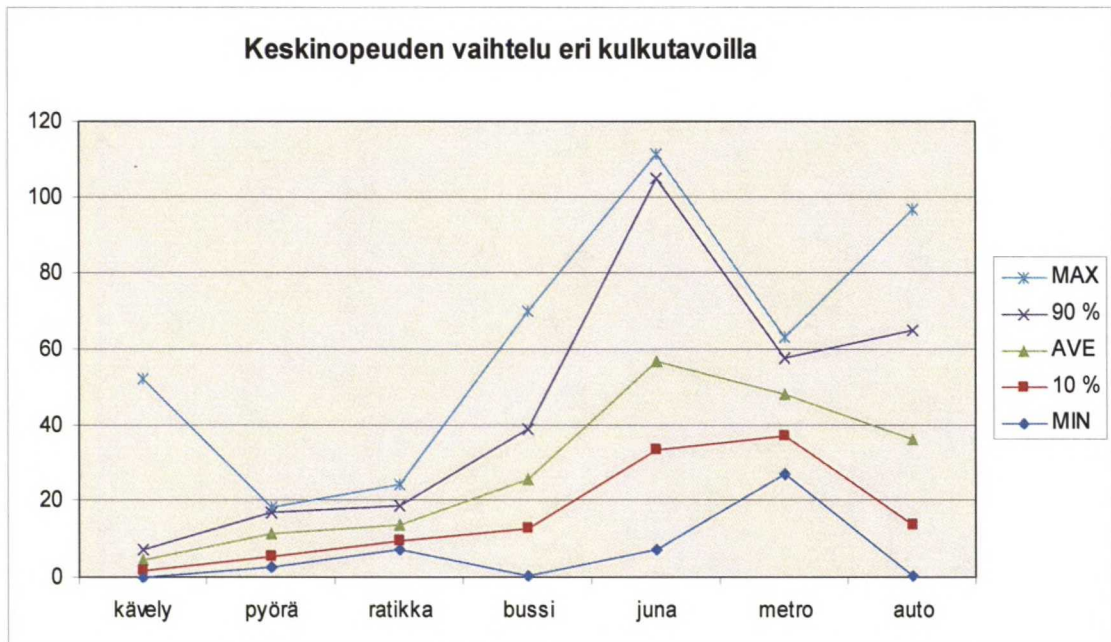
AVE = keskiarvo

10 % = 10. persentiili (10 % arvoista on tätä pienempiä)

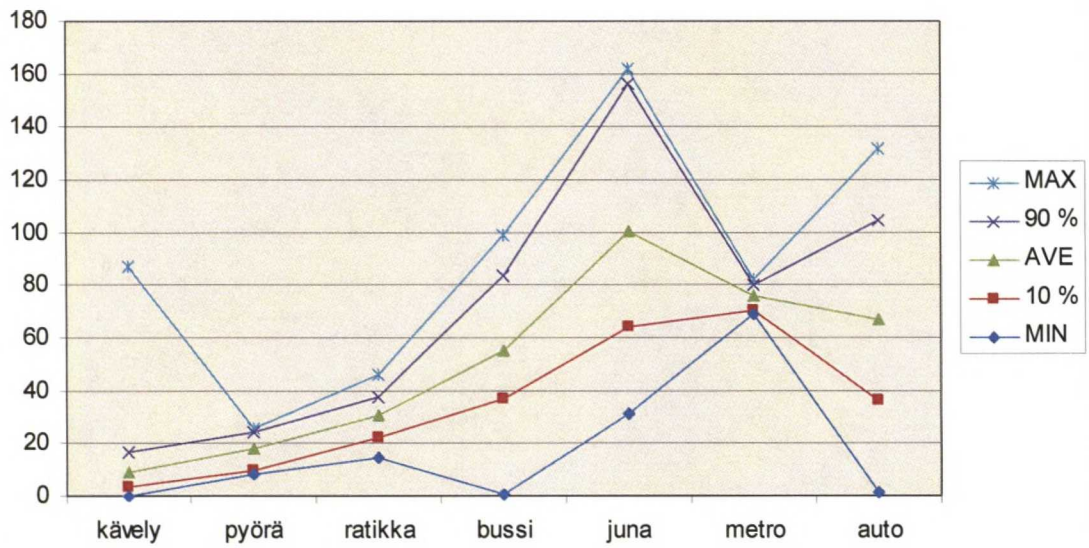
MIN = pienin arvo aineistossa

Kuvia on tulkittu niin, että on epätodennäköistä, että mennään suurimman arvon yli tai jäädään pienimmän alle. 10 %:n ja 90 %:n viivojen välissä ollaan melko todennäköisellä alueella. Jos kaikki viisi lukua ovat hyvin lähellä toisiaan, muuttujan vaihtelu kyseisellä kulkutavalla on pientä, ja muuttujaa voidaan luultavasti käyttää kyseisen kulkutavan tunnistamiseen erityisen hyvin.

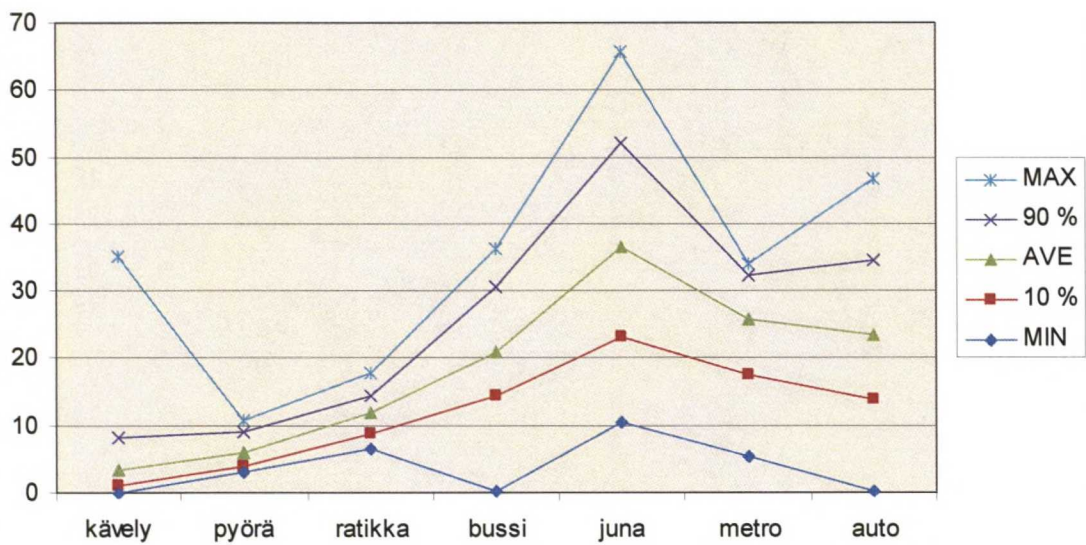
Joistakin kulkutavoista on melko vähän havaintoja, ja kävelyn tunnusluvuissa on selvästi ongelmia (kävelymatkan keskinopeus ei missään tapauksessa voi olla 50 km/h), joten tuloksiin on suhtauduttu pienellä varauksella ja intuitiota on käytetty lisänä algoritmeja laadittaessa.



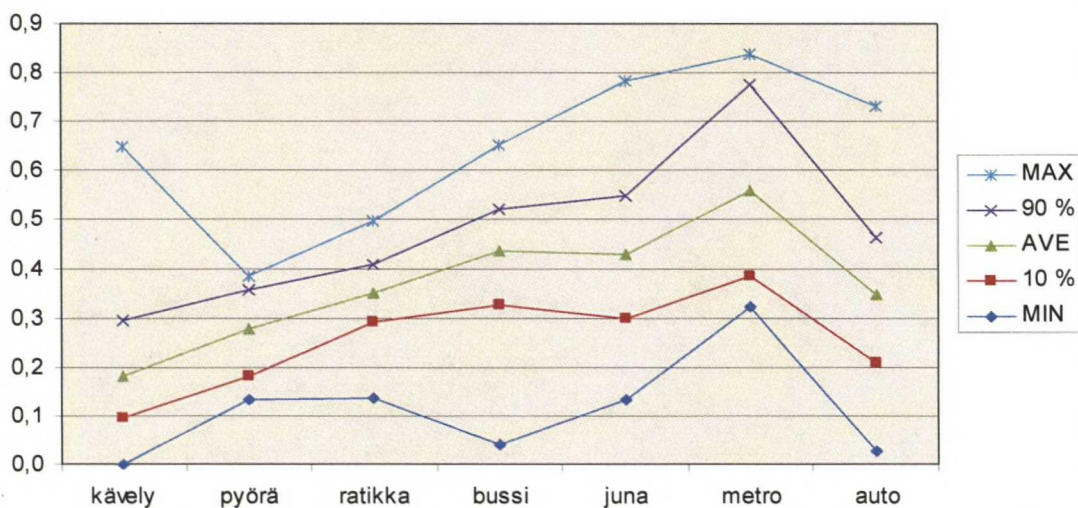
Nopeuden 90. persentiiliin vaihtelu eri kulkutavoilla



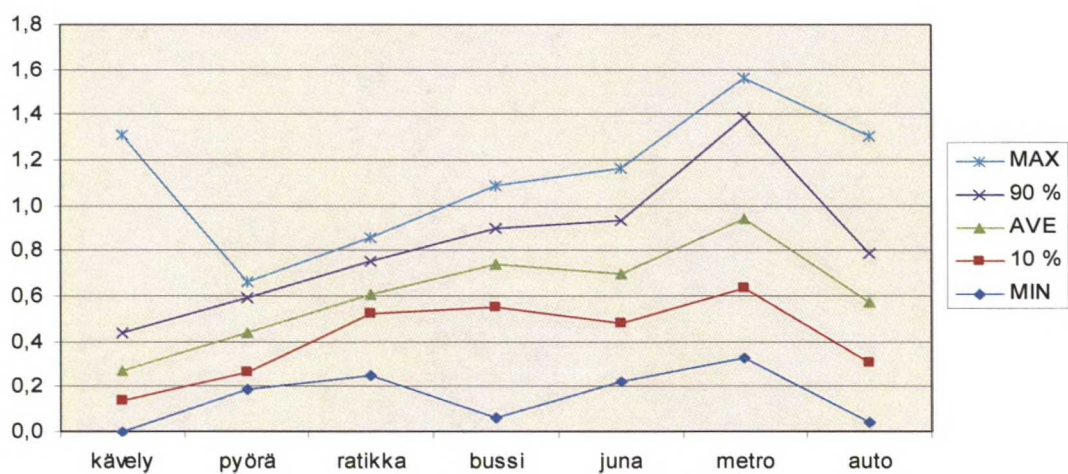
Nopeuden keskihajonnan vaihtelu eri kulkutavoilla



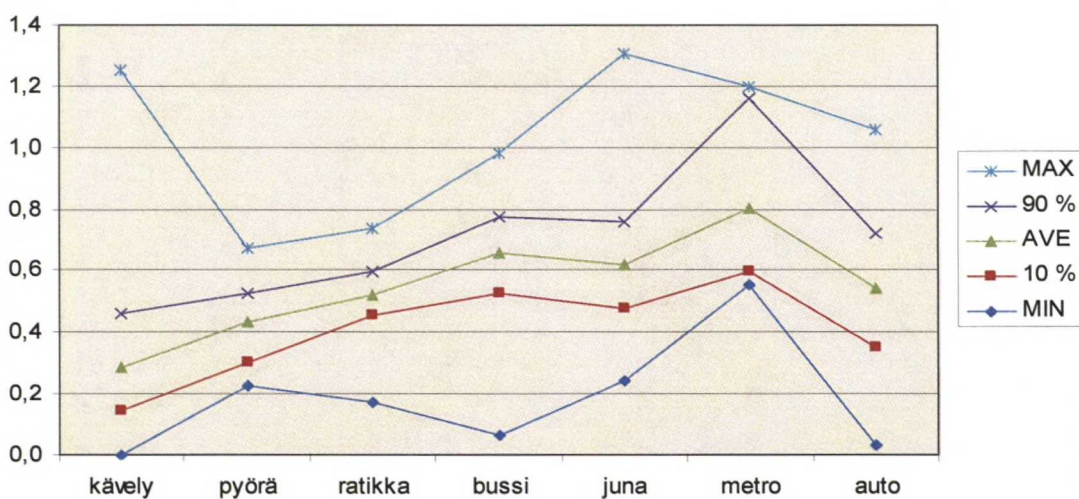
Kiihtyvyyden itseisarvon keskiarvon vaihtelu eri kulkutavoilla



Kiihtyvyyden 90. persenttiin vaihtelu eri kulkutavoilla



Kiihtyvyyden keskihajonnan vaihtelu eri kulkutavoilla



Liite 8. Aineistonjakamisalgoritmin luonnos, vielä kokeiluvaiheessa

Tässä esitetty algoritmi soveltuu ainoastaan aineistolle, joka on tallennettu sekunnin välein.

1) Etsitään kohdat, joissa ihminen on pysähtynyt, vaikka vain hetkeksi (kriteerinä nopeus alle 1,5 km/h 3 sekunnin (peräkkäisen pisteen) ajan, tai pisteitä ei ole tallentunut ollenkaan 10 sekuntiin), ja katkaistaan jokaisesta tällaisesta kohdasta -> saadaan yhtenäinen koordinaattijono pilkottua pienemmiksi jaksoiksi.
2a) Tarkastellaan näin syntyneiden jaksoiden linnuntietä kuljettua matkaa, kestoa ja keskimääräistä linnuntienopeutta (=etäisyys alkupisteestä loppupisteeseen / kesto) <ul style="list-style-type: none">- jos nopeus on alle 8 km/h, arvataan että kyseessä on kävelyjakso- jos nopeus on yli 8 km/h, tarkastellaan, osuvatko matkan alku ja/tai loppu lähelle joukkoliikenteen pysäkkejä<ul style="list-style-type: none">- jos kyllä, niin arvataan kulkutavaksi kyseinen joukkoliikenneväline- jos ei, niin arvataan kulkutavaksi polkupyörä (keskinopeus alle 20 km/h) tai auto Lisäksi, jos edellisen jakson kulkutapa on sama, etäisyys loppupisteeseen on riittävän pieni ja aika jaksoiden välillä riittävän lyhyt (esimerkiksi alle 10 s ja alle 10 m, voi riippua myös kulkutavasta), yhdistetään jaksot edellisen jakson perään. (Tällä menettelyllä saadaan melko hyvin poistettua myös aktiviteettien aikana rakennusten kohdalle syntyvä siksak, koska niistä syntyvien jaksoiden matkat, kestot ja linnuntienopeudet ovat usein järjettömiä suhteessa toisiinsa; järjettömiä jaksoja voidaan poistaa useassa kohdassa algoritmin toteutusta.)
2b) Jotta joukkoliikennevälineiden pysähtyminen asemalle ei aiheuta väärää katkaisua, yhdistetään toisiinsa ne jaksot, joiden <ul style="list-style-type: none">- nopeus on kävelyä suurempi, ja- välinen pysähdys tapahtuu lähellä pysäkkiä ja on riittävän lyhyt (riippuu kulkutavasta), jotta matkustaja todennäköisesti ei ole vaihtanut välinettä
3) Yhdistetään peräkkäiset osuudet, joiden kulkutapa-arvaus on sama ja välinen pysähdys niin lyhyt, että aktiviteettia ei liene tapahtunut
4) Koska edellä kuvattu lähinnä nopeuteen perustuva menettely väistämättä arvaa usein väärin, korjataan ”selvästi” vääriä arvauksia, kun jaksoiden välissä on vain muutama sekunti eli välinettä ei ole ehtinyt vaihtaa, esim. juna-auto-juna -> juna bussi-auto/pyörä-bussi -> bussi auto-pyörä tai pyörä-auto -> auto
5) Poistetaan välistä nopeuden, matkan pituuden tai näiden yhdistelmän perusteella järjettömän tuntuksia jaksoja, jotka todennäköisesti eivät ole oikeita matkoja
6) Analysoidaan näin muodostetut jaksot varsinaisella kulkutavantunnistusalgoritmillä
7) Toistetaan tarvittaessa kohtia 3-6

Liite 9. Tunnistustuloksia pisteytysjärjestelmästä

Tulokset kulkutavanpäättelyalgoritmin versiosta 2.1. Käytetty pisteytysjärjestelmää, jossa muuttujina keskinopeus, nopeuden 90. persentiili ja nopeuden keskihajonta, joukkoliikenteen pysäkit ja matkan aikana tehdyt pysähdykset, keskimääräinen kiihtyvyyden itseisarvo, kiihtyvyyden 90. persentiili ja kiihtyvyyden keskihajonta.

Oikeat matkat		Algoritmin löytämät kulkutavat								
Matkan osia		ei								onnistumis- prosentti
Kulkutapa	aineistossa	tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	529	0	466	20	16	17	5	4	1	88,1 %
pyörä	24	0	5	17	0	2	0	0	0	70,8 %
auto	151	0	5	2	104	32	0	7	1	68,9 %
bussi	146	0	2	1	13	130	0	0	0	89,0 %
raitiovaunu	75	0	0	1	2	15	57	0	0	76,0 %
juna	34	0	1	0	1	2	0	29	1	85,3 %
metro	22	0	1	0	4	3	0	0	14	63,6 %
yhteensä	981	0	480	41	140	201	62	40	17	
Oikein yhteensä: 83,3 %										

Tulokset kulkutavanpäättelyalgoritmin versiosta 2.2. Käytetty pisteytysjärjestelmää, jossa muuttujina keskinopeus, nopeuden 90. persentiili ja nopeuden keskihajonta, joukkoliikenteen pysäkit, keskimääräinen kiihtyvyyden itseisarvo, kiihtyvyyden 90. persentiili ja kiihtyvyyden keskihajonta.

Oikeat matkat		Algoritmin löytämät kulkutavat								
	Matkan osia	ei								onnistumis- prosentti
Kulkutapa	aineistossa	tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	529	0	466	20	12	20	8	2	1	88,1 %
pyörä	24	0	5	16	0	1	2	0	0	66,7 %
auto	151	0	5	2	106	32	0	3	3	70,2 %
bussi	146	0	2	0	18	125	1	0	0	85,6 %
raitiovaunu	75	0	0	3	4	17	51	0	0	68,0 %
juna	34	0	1	0	1	5	0	26	1	76,5 %
metro	22	0	1	0	5	4	0	0	12	54,5 %
yhteensä	981	0	480	41	146	204	62	31	17	
Oikein yhteensä: 81,8 %										

Tulokset kulkutavanpäättelyalgoritmin versiosta 3.0. Käytetty pisteytysjärjestelmää, jossa muuttujina keskinopeus, nopeuden 90. persentiili ja nopeuden keskihajonta.

Oikeat matkat		Algoritmin löytämät kulkutavat								
	Matkan osia	ei								onnistumis- prosentti
Kulkutapa	aineistossa	tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	529	0	480	5	31	0	12	1	0	90,7 %
pyörä	24	0	5	18	0	0	1	0	0	75,0 %
auto	151	0	5	2	129	0	1	14	0	85,4 %
bussi	146	0	3	0	139	0	2	2	0	0,0 %
raitiovaunu	75	0	3	11	39	0	22	0	0	29,3 %
juna	34	0	0	0	16	0	1	17	0	50,0 %
metro	22	0	1	0	21	0	0	0	0	0,0 %
yhteensä	981	0	497	36	375	0	39	34	0	
Oikein yhteensä: 67,9 %										

Tulokset kulkutavanpäättelyalgoritmin versiosta 3.1. Käytetty pisteytysjärjestelmää, jossa muuttujina keskinopeus, nopeuden 90. persentiili ja nopeuden keskihajonta, keskimääräinen kiihtyvyyden itseisarvo, kiihtyvyyden 90. persentiili ja kiihtyvyyden keskihajonta.

Oikeat matkat		Algoritmin löytämät kulkutavat								onnistumis- prosentti
Kulkutapa	Matkan osia aineistossa	ei tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	529	0	466	17	31	4	11	0	0	88,1 %
pyörä	24	0	5	15	0	0	4	0	0	62,5 %
auto	151	0	5	1	128	5	1	8	3	84,8 %
bussi	146	0	2	0	112	30	1	1	0	20,5 %
raitiovaunu	75	0	0	7	40	0	28	0	0	37,3 %
juna	34	0	1	0	11	4	0	16	2	47,1 %
metro	22	0	1	0	11	5	0	0	5	22,7 %
yhteensä	981	0	480	40	333	48	45	25	10	
Oikein yhteensä: 70,1 %										

Liite 10. Tulokset testistä, jossa poimittiin GPS-aineistosta joka 5. tai joka 10. tietue

Tulokset algoritmin versiolla 1.11, kun simuloitiin 5 sekunnin välein tallennettua GPS-aineistoa.

Oikeat matkat		Algoritmin löytämät kulkutavat								
	Matkan osia	ei								
Kulkutapa	aineistossa	tunnis-	kävely	pyörä	auto	bussi	raitio-	juna	metro	onnistumis-
		tettu					vaunu			prosentti
kävely	529	0	482	3	18	21	4	1	0	91,1 %
pyörä	24	0	5	18	1	0	0	0	0	75,0 %
auto	151	0	5	2	100	38	0	6	0	66,2 %
bussi	146	0	4	0	14	115	9	4	0	78,8 %
raitiovaunu	75	0	3	1	3	7	61	0	0	81,3 %
juna	34	0	1	0	6	2	0	25	0	73,5 %
metro	22	0	1	0	13	8	0	0	0	0,0 %
yhteensä	981	0	501	24	155	191	74	36	0	
Oikein yhteensä: 81,7%										

Tulokset algoritmin versiolla 1.11, kun simuloitiin 10 sekunnin välein tallennettua GPS-aineistoa.

Oikeat matkat		Algoritmin löytämät kulkutavat								
Matkan osia		ei								onnistumisprosentti
Kulkutapa	aineistossa	tunnistettu	kävely	pyörä	auto	bussi	raitiovaunu	juna	metro	
kävely	529	0	479	6	19	21	4	0	0	90,5 %
pyörä	24	0	5	18	1	0	0	0	0	75,0 %
auto	151	0	6	1	99	39	0	6	0	65,6 %
bussi	146	0	3	1	14	114	9	5	0	78,1 %
raitiovaunu	75	0	2	1	4	6	62	0	0	82,7 %
juna	34	0	0	0	7	2	0	25	0	73,5 %
metro	22	0	1	0	15	6	0	0	0	0,0 %
yhteensä	981	0	496	27	159	188	75	36	0	
Oikein yhteensä: 81,2 %										

Tulokset algoritmin versiolla 2.0, kun simuloitiin 5 sekunnin välein tallennettua GPS-aineistoa.

Oikeat matkat		Algoritmin löytämät kulkutavat								
Matkan osia		ei								onnistumisprosentti
Kulkutapa	aineistossa	tunnistettu	kävely	pyörä	auto	bussi	raitiovaunu	juna	metro	
kävely	529	0	479	11	23	6	6	3	1	90,5 %
pyörä	24	0	5	18	0	1	0	0	0	75,0 %
auto	151	0	5	2	112	28	1	3	0	74,2 %
bussi	146	0	3	0	30	112	1	0	0	76,7 %
raitiovaunu	75	0	1	0	1	15	58	0	0	77,3 %
juna	34	0	0	0	3	5	0	26	0	76,5 %
metro	22	0	1	0	8	3	0	0	10	45,5 %
yhteensä	981	0	494	31	177	170	66	32	11	
Oikein yhteensä: 83,1 %										

Tulokset algoritmin versiolla 2.0, kun simuloitiin 10 sekunnin välein tallennettua GPS-aineistoa.

Oikeat matkat		Algoritmin löytämät kulkutavat								onnistumis- prosentti
Kulkutapa	Matkan osia aineistossa	ei tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	529	0	477	17	22	4	6	2	1	90,2 %
pyörä	24	0	5	19	0	0	0	0	0	79,2 %
auto	151	0	6	2	117	22	1	3	0	77,5 %
bussi	146	0	2	3	58	82	1	0	0	56,2 %
raitiovaunu	75	0	1	5	12	10	47	0	0	62,7 %
juna	34	0	0	0	3	5	0	26	0	76,5 %
metro	22	0	1	0	10	2	0	0	9	40,9 %
yhteensä	981	0	492	46	222	125	55	31	10	
Oikein yhteensä: 79,2 %										

Liite 11. Tulokset validointiaineistolla tehdyistä testeistä

Tulokset algoritmin versiolla 1.11.

Oikeat matkat		Algoritmin löytämät kulkutavat								onnistumis- prosentti
Kulkutapa	Matkan osia aineistossa	ei tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	16	0	16	0	0	0	0	0	0	100,0 %
pyörä	3	0	1	1	1	0	0	0	0	33,3 %
auto	2	0	0	0	1	1	0	0	0	50,0 %
bussi	8	0	0	0	0	8	0	0	0	100,0 %
raitiovaunu	0	0	0	0	0	0	0	0	0	0,0 %
juna	1	0	0	0	0	1	0	0	0	0,0 %
metro	0	0	0	0	0	0	0	0	0	0,0 %
yhteensä	30	0	17	1	2	10	0	0	0	
Oikein yhteensä: 86,7 %										

Tulokset algoritmin versiolla 2.0.

Oikeat matkat		Algoritmin löytämät kulkutavat								onnistumis- prosentti
Kulkutapa	Matkan osia aineistossa	ei tunnis- tettu	kävely	pyörä	auto	bussi	raitio- vaunu	juna	metro	
kävely	16	0	16	0	0	0	0	0	0	100,0 %
pyörä	3	0	1	2	0	0	0	0	0	66,7 %
auto	2	0	0	0	1	1	0	0	0	50,0 %
bussi	8	0	0	0	0	8	0	0	0	100,0 %
raitiovaunu	0	0	0	0	0	0	0	0	0	0,0 %
juna	1	0	0	0	0	1	0	0	0	0,0 %
metro	0	0	0	0	0	0	0	0	0	0,0 %
yhteensä	30	0	17	2	1	10	0	0	0	
Oikein yhteensä: 90,0 %										